



Europäisches Patentamt
European Patent Office
Office européen des brevets



(11) Publication number:

0 159 884 B1

(12)

EUROPEAN PATENT SPECIFICATION

(45) Date of publication of patent specification: 10.02.93 (51) Int. Cl.⁵: C12N 15/82, C12P 21/02

(21) Application number: 85302593.0

(22) Date of filing: 12.04.85

(54) Heat shock promoter and gene.

(30) Priority: 13.04.84 US 599993

(43) Date of publication of application:
30.10.85 Bulletin 85/44

(49) Publication of the grant of the patent:
10.02.93 Bulletin 93/06

(84) Designated Contracting States:
AT BE CH DE FR GB IT LI LU NL SE

(56) References cited:
EP-A- 0 116 718

JOURNAL OF MOLECULAR AND APPLIED GENETICS, vol. 1, no. 4, 1982, pages 301-314, Raven Press, New York, US; F. SCHÖFFL et al.: "An analysis of mRNAs for a group of heat shock proteins of soybean using cloned cDNAs"

PLANT MOLECULAR BIOLOGY, vol. 2, 1983, pages 269-278, Martinus Nijhoff/Dr W. Junk Publishers, The Hague, NL; F. SCHÖFFL et al.: "Identification of a multigene family for small heat shock proteins in soybean and physical characterization of one individual gene coding region"

(73) Proprietor: LUBRIZOL GENETICS INC.
29400 Lakeland Boulevard
Wickliffe Ohio 44092(US)

Proprietor: THE UNIVERSITY OF GEORGIA RESEARCH FOUNDATION, INC.
The University of Georgia, 612 Boyd
Graduate Studies Research Center
Athens, Georgia 30602(US)

Proprietor: THE UNIVERSITY OF FLORIDA
223 Grinter Hall
Gainesville, Florida 32611(US)

(72) Inventor: Key, Joe L.
3068 10th Street
Boulder Colorado 80020(US)
Inventor: Gurley, William B.
2307 Southwest 32nd Place No. 5
Gainesville Florida 32608(US)
Inventor: Nagao, Ronald T.
185 Brookwood Drive
Athens Georgia 30605(US)
Inventor: Schoeffl, Friedrich
Im Bergsiek 35
W-4800 Bielefeld 15(DE)

Note: Within nine months from the publication of the mention of the grant of the European patent, any person may give notice to the European Patent Office of opposition to the European patent granted. Notice of opposition shall be filed in a written reasoned statement. It shall not be deemed to have been filed until the opposition fee has been paid (Art. 99(1) European patent convention).

THE GENETIC MANIPULATION OF PLANTS AND ITS APPLICATION TO AGRICULTURE, edited by G.R. Stewart et al., 1984, pages 129-140; F. SCHÖFFL et al.: "Soybean heat shock proteins: temperature regulated gene expression and the development of thermotolerance"

NATURE, vol.303, 19th May 1983, pages 209-213, Macmillan Journals Ltd; L. HERRERA-ESTRELLA et al.: "Expression of chimaeric genes transferred into plant cells using a Ti-plasmid-derived vector"

CHEMICAL ABSTRACTS, vol. 105, 1986, page 192, abstract no. 128581u, Columbus, Ohio, US; J. SCHELL et al.: "Transfer and regulation of expression of chimeric genes in plants", & COLD SPRING HARBOR SYMPOSIUM ON QUANT. BIOL. 1985, 50(MOL. BIOL. DEV.), 421-31

SCIENCE, vol. 223, February 1984, pages 496-498; R.B. HORSCH et al.: "Inheritance of functional foreign genes in plants"

CHEMICAL ABSTRACTS, vol. 100, 1984, page 142, abstract no. 62725g, Columbus, Ohio, US; J.T. LIS et al.: "New heat shock puffs and beta-galactosidase activity resulting from transformation of *Drosophila* with an hsp70-lacZ hybrid gene", & CELL (CAMBRIDGE, MASS.) 1983, 35(2, pt. 1), 403-10

CHEMICAL ABSTRACTS, vol. 79, 1973, page 193, abstract no. 123826n, 1973, Columbus, Ohio, US; C.B. JOHNSON et al.: "Expression of lambda-plac5 DNA in cultured cells of a higher plant", & NATURE (LONDON), NEW BIOL. 1973, 244(134), 105-7

BIOLOGICAL ABSTRACTS/RRM, no. 29054462, Journal of cellular biochemistry supplement (US), 1985, no. 9, part C, page 254, Biosciences Information Services, US; W.B. GURLEY et al.: "Heterologous expression of a soybean gene encoding a small heat-shock protein in sunflower tumors"

Inventor: Czarnecka, Eva
2307 Southwest 32nd Place No. 5
Gainesville Florida 32608(US)

Ⓐ Representative: Fisher, Adrian John et al
CARPMAELS & RANSFORD 43 Bloomsbury
Square
London WC1A 2RA(GB)

Description

A class of genes known as heat shock or stress genes occurs in all organisms from bacteria to man. Transcription of these genes is initiated following a stress treatment (e.g., heat shock) and translation of the transcripts produces proteins that probably protect the cell temporarily. During stress, normal polyribosomes quickly break down to monoribosomes which are then used to translate the heat shock mRNA's. Normal mRNA's present before the stress treatment are in some way protected during the stress period and they can be re-used in translation following termination of the stress. The production of heat shock mRNA's and proteins is only a temporary phenomenon and the expression of the heat shock genes levels off after a few hours and then declines. If the temperature is increased slowly rather than in a single step, an organism can withstand temperatures which would otherwise be lethal, i.e., the organism can adapt to higher temperatures.

Background of the Invention

Over two decades ago, it was discovered (Ritossa, F. (1962) *Experientia* 18:571-573) that specific puffing patterns in the polytene chromosomes of *Drosophila busckii* could be induced by a brief heat shock. The puffing positions of *Drosophila* species polytene chromosomes are positions where there is active synthesis of mRNA, thus indicating active gene loci (Beerman, W. (1956) Cold Spring Harbor Symp. Quant. Biol. 21:217-232). Since then it has been shown that a variety of agents, e.g., arsenite or anaerobic conditions, can induce responses similar to those induced by heat, suggesting that a more appropriate name for these genes should be "stress genes". However, the nomenclature of "heat shock" genes is now well established and will be used in the remainder of this application.

Since the early 1950's it was known that the pattern of puffs in the polytene chromosomes of Dipteran larvae changed in a regular manner during development, and it was shown that these changes were controlled by ecdysteroid hormones (Clever, U. and P. Karlson (1960) *Exp. Cell. Res.* 20:623-626; Becker, H.-J. (1962) *Chromosoma* 13:341-384). In particular, it was shown that the pattern of puffing was disrupted by a brief heat shock (or treatment with certain chemicals) and resulted in the appearance of three new puffs. The induction of these new puffs was very rapid and occurred within minutes of the heat shock treatment, but the induction was transient. For example, when the temperature was raised from 25° → 37° C, the puffs reached maximum size within 30 minutes and then regressed. At the same time, all puffs active before the heat shock regressed after the treatment. The heat shock puffing response was also found to occur in all tissues studied and at all stages of development.

It was later found that the heat shock treatment induced synthesis of a small number of polypeptides and repressed the synthesis of most others (Tissieres, A. et al. (1974) *J. Mol. Biol.* 84:389-398) and the mRNA produced at the new heat shock induced puffs was shown to code for the newly induced polypeptides (Lewis, M. J. et al. (1975) *Proc. Nat. Acad. Sci. USA* 72:3604-3608).

Within a few minutes of heat shock, all polyribosomes break down and are quickly replaced by a new polyribosome peak which contains heat-shock protein mRNA. This mRNA has been hybridized back to the heat shock induced puffs and has been translated in vitro into heat shock proteins (McKenzie, S. et al. (1977) *J. Mol. Biol.* 117:279-283; Mirault, M. E. et al. (1978) Cold Spring Harb. Symp. Quant. Biol. 42:819-827). It is of interest to note that even though all polyribosomes break down and the newly induced hsmRNA's are selectively translated, most normal mRNA's persist during heat shock (Ashburner, N. and J. F. Bonner (1979) *Cell* 17:241ff.).

Most of the early work on heat shock genes was done with *Drosophila* species. However, in 1978, analogous stress responses were found in chick embryonic fibroblasts (Kelly, P. and M. J. Schlesinger (1978) *Cell* 15:1277-1286), in Chinese hamster ovary cells (Bouche, G. et al. (1979) *Nucleic Acids Research* 7:1739-1747), in *Escherichia coli* (Lemeaux, P. G. et al. (1978) *Cell* 13:427-434), in yeast (Miller, M. J. et al. (1979) *Proc. Nat. Acad. Sci. USA* 76:5222-5225), in *Naegleria* (Walsh, C. (1980) *J. Biol. Chem.* 255:2626-2632), in *Tetrahymena* (Fink, K. and E. Zeuthen (1978) *ICN-UCLA Symp. Mol. Cell. Biol.* 12:103-115) and in many other species, including plants (Barnett, T. et al. (1980) *Dev. Genet.* 1:331-340). A similar pattern of heat shock protein synthesis has also been reported for tobacco and soybean cells growing in solution culture (Barnett, T., et al. (1980) *supra*) as that reported for soybean seedling tissue. It was also shown that the effects of trauma on vertebrate cells was similar to the effects of heat shock (Hightower, L. E. and F. P. Whit (1981) *J. Cell. Physiol.* 108:261).

The transcriptional and translational control of heat shock genes may be autoregulatory. Thus the activity of these genes may be controlled by the concentrations of the heat shock proteins present in the cells. Therefore, inducers of heat shock genes would be factors that either destroyed the heat shock

proteins or rendered them to be effectively unavailable within the cell, e.g., by binding to various cell organelles.

The activation and subsequent repression of heat shock genes in *Drosophila* has been studied by the introduction of cloned segments into *Drosophila* cells. The P-element-mediated transformation system, which permits introduction of cloned *Drosophila* genes into the *Drosophila* germline, was used (Rubin, G. M. and A. C. Spradling (1982) *Science* 218:348-353). A gene integrated in this way is often present as a stable, single copy and has a relatively constant activity at a variety of chromosomal locations (Scholnick, S. B. et al. (1983) *Cell* 34:37-45; Goldberg, D. et al. (1983) *Cell* 34:59-73; Spradling, A. C. and G. M. Rubin (1983) *Cell* 34:47-57). In particular, the *Drosophila* hsp70 gene was fused in phase to the *E. coli* β -galactosidase structural gene, thus allowing the activity of the hybrid gene to be distinguished from the five resident hsp70 heat shock genes in the recipient *Drosophila*. *Drosophila* heat shock genes have also been introduced and their activity studied in a variety of heterologous systems, and, in particular, in monkey COS cells (Pelham, H. R. B. (1982) *Cell* 30:517-528; Mirault, M.-E. et al. (1982) *EMBO J.* 1:1279-1285) and mouse cells (Corces, V. et al. (1981) *Proc. Nat. Acad. Sci.* 78:7038-7042).

The hybrid hsp70-lacZ gene appeared to be under normal heat shock regulation when integrated into the *Drosophila* germ line (Lis, J. T. et al. (1983) *Cell* 35:403-410). Three different sites of integration formed large puffs in response to heat shock. The kinetics of puff formation and regression were exactly the same as those of the 87C locus, the site from which the integrated copy of the hsp70 gene was isolated. The insertion of the 7 kilobase *E. coli* β -galactosidase DNA fragment into the middle of the hsp70 structural gene appeared to have had no adverse effect on the puffing response. The β -galactosidase activity in the transformants was regulated by heat shock.

Deletion analysis of the *Drosophila* hsp70 heat shock promoter has identified a sequence upstream from the TATA box which is required for heat shock induction. This sequence contains homology to the analogous sequence in other heat shock genes and a consensus sequence CTxGAAxxTTCxAG has been constructed (Pelham, H. R. B. and M. Bienz (1982) *EMBO J.* 1:1473-1477). When synthetic oligonucleotides, whose sequence was based on that of the consensus sequence, were constructed and placed upstream of the TATA box of the herpes virus thymidine kinase gene (tk) (in place of the normal upstream promoter element), then the resultant recombinant genes were heat-inducible both in monkey COS cells and in *Xenopus* oocytes. The tk itself is not heat inducible and probably no evolutionary pressure has occurred to make it heat inducible, but the facts above indicate that tk can be induced by a heat shock simply by replacing the normal upstream promoter element with a short synthetic sequence which has homology to a heat shock gene promoter.

An inverted repeat sequence upstream of the TATA box is a common feature of many of the heat shock promoters which have been studied (Holmgren, R. et al. (1981) *Proc. Nat. Acad. Sci. USA* 78:3775-3778). In five of the seven *Drosophila* promoters, this inverted repeat is centered at the 5'-side of the penultimate A residue of the consensus sequence, but the sequence of the inverted repeat itself is not conserved (Pelham, H. R. B. (1982) *Cell* 30:517-528). In some cases, however, the inverted repeat sequence occurs upstream from the TATA box and the consensus sequence is not present. In these cases there is no heat inducibility so the presence of the inverted repeat does not substitute for the consensus sequence.

The functional significance of the heat shock response is not known. Presumably it functions to protect the cell against environmental stress and to allow the cell to continue its function after the stress situation has passed. These conclusions are supported by a phenomenon known as "acquired thermotolerance". Cells exposed to a single heat shock, or some other stress, are relatively protected against the effects of a second, otherwise lethal heat shock (Li, G. C. and G. M. Hahn (1978) *Nature* 274:699-701; Henle, K. J. and L. A. Dethlefsen (1978) *Cancer Res.* 38:1843-1851; Mitchell, H. K. et al. (1979) *Dev. Genet.* 1:181-192; McAlister, L. and D. B. Finkelstein (1980) *Biochem. Biophys. Res. Commun.* 93:819-824).

In higher plants, the heat shock (hs) phenomenon was first discovered at the level of protein synthesis in soybeans (Key, J. L. et al. (1981) *Proc. Nat. Acad. Sci. USA* 78:3526-3530; Barnett, T. et al. (1980) *supra*). A number of other plants, e.g., pea, millet, corn, sunflower, cotton and wheat, respond similarly to soybean in that a large number of new proteins of similar molecular weight are induced by a heat shock treatment. The major differences that occur among species are the optimum temperature of induction of hs-proteins, the breakpoint temperature (i.e., above this temperature is lethal), the distribution of the 15-20 kD heat shock proteins on two-dimensional gels and the relative level of normal protein synthesis that occurs during heat shock. It has been shown that an elevation of temperature from 28°C to 40°C induced *de novo* synthesis of several major groups of hs-proteins (hsp) whose molecular weights resemble those found for *Drosophila*. However, there is a marked difference in the complexity of the low molecular weight (lmw) group of hsp's between these two organisms. *Drosophila* synthesizes four hsp's of 22, 23, 26 and 27 kilodaltons; soybean produces more than 20 hsp's in the molecular weight range of 15-18 kilodaltons.

- The translational preference for hs-mRNA's, while marked, appeared less pronounced in the soybean system (Key, J. L. et al. (1981) *supra*) than in *Drosophila* (Storti, R. V. et al. (1980) *Cell* 22:825-834). The induction of a new set of hs-specific mRNA's in soybean was suggested by *in vitro* translation of poly(A)⁺ RNA. Additional evidence for the existence of novel RNA in heat stressed plants was provided by sucrose gradient analysis which showed the accumulation of a 0.49×10^6 dalton RNA during hs of tobacco and cowpea leaves (Dawson, W. O. and G. L. Grantham (1981) *Biochem. Biophys. Res. Commun.* 100:23-30). In *Drosophila*, where transcriptional control of hs protein synthesis is evident, attempts have been made to find signal structures for coordinate expression of these genes (Holmgren, R. et al. (1981) *Proc. Nat. Acad. Sci. USA* 78:3775-3778). The influence of hs on poly(A)⁺ mRNA's of soybean has been assessed using cDNA/poly(A)⁺ RNA hybridization and cloned cDNA/northern blot hybridization analyses (Schoffl, F. and J. L. Key (1982) *J. Mol. Appl. Genet.* 1:301-314). The hs response in soybean is characterized by the appearance of a new highly abundant class of poly(A)⁺ RNA's consisting of some twenty different sequences of an average length of 800 to 900 nucleotides and a decrease in total poly(A)⁺ RNA complexity associated with changes in relative abundance of the 28°C sequences. The poly(A)⁺ RNA's of this new abundant class are present at some 15,000 to 20,000 copies per cell after 2 hours of hs at 40°C. The genes for these four *Drosophila* hsp's comprise a small hs-gene family with similar sequences which are also related to that of α -crystallin (Ingolia, T. D. and E. A. Craig (1982) *Proc. Nat. Acad. Sci. USA* 79:2360-2364) implying that certain structural domains (possibly for functional aggregation) are shared by these proteins. The *lmw*-hsp genes in soybeans are the most actively expressed and coordinately regulated genes under hs conditions (Schoffl, F. and J. L. Key (1982) *J. Mol. Appl. Genet.* 1:301-314). Their hsp's are commonly associated with purified nuclei at high temperature, however, and disaggregate at low temperature (Key, J. L. et al. (1982) In: Schlesinger, M. J., Ashburner, M. and A. Tissieres (eds.) *Heat shock, from Bacteria to Man*. Cold Spring Harbor Laboratory, pp. 329-336). This indicates a common function for these proteins in hs-response which is possibly related to common structural features in proteins and genes. The *lmw*-hsp genes are subdivided into eight classes defined by sequence homologies among poly(A)⁺ mRNA's. Two of the eight classes are particularly interesting with respect to gene expression, because they represent the extreme components of the *lmw*-hsp genes. These are designated classes I and II; I consists of 13 closely related hsp's genes, while II comprises only 1hsp which has no known sequence homology to other hs-genes. Later information showed that class II could be grouped with class I. The separation into the two classes was originally made on the basis of a probe distal to the 3'-translated end of pE2019.
- A wide range of crop plants respond to elevated temperatures of heat shock conditions by synthesizing a large number (30 or more) of hs-proteins (Key, J. L. et al. (1983) *Current Topics In Plant Biochemistry and Physiology*, eds. D. D. Randall, D. G. Blevins, A. L. Larson and B. J. Rapp. Vol. 2, Univ. of Missouri, Columbia, pp. 107-117). The high molecular weight hs-proteins were electrophoretically similar among the species. The more complex pattern of low molecular weight (15-27 kd) hs-proteins showed much more electrophoretic heterogeneity between species. Certainly a given soybean hs-cDNA clone showed greater cross hybridization to different soybean hs-poly(A) RNA's than to any hs-RNA from other species, and this limited hybridization with other species was consistent with the observed electrophoretic heterogeneity of the low molecular weight hs-proteins.
- The evolutionary conservation of the hs-response across the spectrum of organisms from bacteria to man suggests an essential function(s) for the hs-proteins. Empirically, one function is to provide thermal protection or thermotolerance to otherwise non-permissive hs temperature (Schlesinger, M. et al. (1982) *Heat shock from bacteria to man*. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY p.329). Apparently those hs-proteins which are synthesized at a permissive heat shock temperature allow organisms to continue the synthesis of hs-proteins and hs-mRNA's at still higher temperatures and to survive what would be normally lethal temperatures (Key, J. L. et al. (1982) In: *Heat Shock from Bacteria to Man*. M. J. Schlesinger, M. Ashburner and A. Tissieres, eds. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY p.329). A permissive heat shock temperature is defined here as a temperature which is high enough to induce the heat shock response but not high enough to be lethal. Temperatures above break point temperature are lethal for plants which have not acquired a thermotolerance. In soybean the break point temperature is about 40°C. It has previously been shown that soybean seedlings survive incubation at a lethal temperature by prior incubation at a permissive hs-temperature (Key, J. L. et al. (1983) In: *NATO Advanced Studies Workshop on Genome Organization and Expression in Plants*. L. Dure, ed. Plenum Press).
- Several different treatment regimes of permissive heat shock result in the development of thermotolerance in the soybean seedling. These treatments include: (a) a 1- to 2-hour continuous heat shock at 40°C followed by a 45°C incubation; (b) a 30-minute h at shock at 40°C followed by 2 to 3 hours at 28°C prior to the shift to 45°C; (c) a 10-minute heat shock at 45°C followed by about 2 hours at 28°C prior to

the shift to 45°C; and (d) treatment of seedlings with 50µM arsenite at 28°C for 3 hours or more prior to the shift to 45°C. The important feature which these treatments have in common is the induction of synthesis and accumulation of heat shock proteins prior to incubation at the potentially lethal temperature. In fact, it has been shown that both hs-mRNA and hs-protein synthesis do occur at 45°C if the seedlings had earlier been exposed to one of the conditions described above. A likely role(s) for the hs-proteins is to protect vital functions and structures (e.g., transcription, translation and the machinery of energy production) during heat shock and to permit normal functions to return rapidly when favorable temperatures are re-established. It is known that recovery of normal mRNA and protein synthesis occurs rapidly when the temperature is shifted back to normal (e.g., 28°C) (Key, J. L. et al. (1981) *Proc. Nat. Acad. Sci. USA* 78:3526-3530; Schlesinger, M. J. et al. (1982) *Trends Biochem. Sci.* 7:222-225). The resumption of normal protein synthesis utilizes mRNA's conserved during heat shock as well as those newly synthesized during recovery, and there is no detectable synthesis of heat-shock proteins after 3-4 hours at the normal temperature. However, those heat shock proteins that were synthesized during the 40°C heat shock (recognized by the incorporation of ³H-leucine) are very stable during a subsequent chase in non-radioactive leucine, regardless of whether the chase is accomplished at 28°C or 40°C; approximately 80% of the label is retained in the heat shock proteins during a 20-hour chase.

The acquisition of thermotolerance appears to depend not only upon the synthesis of heat shock proteins but also on their selective cellular localization. In soybean seedlings, several hs-proteins become selectively localized in or associated with nuclei, mitochondria and ribosomes in a state that causes them to isolate in gradient-purified fractions of these organelles. Specifically, the complex group of 15-18 kilodalton hs-proteins selectively localize in these fractions during heat shock of soybean seedlings. The selective localization of hs-proteins is temperature dependent. The hs-proteins (except the 22-24 kd. hs-proteins which attach to the mitochondrial fraction) chase from the organelle fractions during a 4-hour incubation at 28°C and they remain organelle associated during a chase at heat shock temperature. In addition, a second heat shock following a 4-hour 28°C chase results in rapid (within 15 minutes) reassociation of hs-proteins with the organelle fractions. This association of heat shock proteins with nuclei could be explained by the hs-proteins becoming "chromatin proteins" or possibly a part of the matrix structure; both suggestions have been offered following localization studies in the *Drosophila* system (Arrigo, A. P. et al. (1980) *Dev. Biol.* 78:86-103). These findings are in basic agreement with autoradiographic results which localized hs-proteins to interband regions of polytene chromosomes (Velazquez, J. et al. (1980) *Cell* 20:679-689 and (1984) *Cell* 36:655-662).

Most of the heat shock work in plants has been done with etiolated seedlings, largely due to ease of manipulation. Heat shock proteins have not been extensively analysed in the green tissues of normal plants, but it has been shown that hs-mRNA's accumulate in green leaf tissue to levels similar to those of etiolated seedlings. Additionally, most experimental work has been done using a large temperature shift of about 10°C. The response to such a non-physiological shift, however, is mimicked both at the level of hs-mRNA and hs-protein synthesis and accumulation, by a gradual increase from 28°C to 47.5°C in the case of soybean. Thus, the results from what may appear to be non-physiological experiments can be duplicated with etiolated seedlings and green plants under more normal physiological conditions of heat shock, which indeed probably occur in the normal plant environment.

Summary of the Invention

Four heat shock genes of soybeans have been cloned and sequenced. The heat shock promoter fragments of these four heat shock genes have been subcloned and have been genetically engineered into a T-DNA shuttle vector. These recombinant DNA fragments, i.e., a vector linked to a T-DNA shuttle vector containing a soybean heat shock gene promoter, can then be transferred with the aid of a helper plasmid into *Agrobacterium tumefaciens* where the recombinant DNA fragment is integrated into the Ti-plasmid. The T-DNA portion of the Ti-plasmid can then be transferred to a plant genome, thus allowing transformation of the plant.

Since the heat shock gene promoter is also transferred to the plant genome and is activated temporarily after a heat shock or stress treatment, it is useful to incorporate foreign genes into the recombinant DNA plasmids in such a position as to be expressible under the control of the heat shock gene promoter. Such incorporated foreign genes can be utilized to recognize plant cells transformed by T-DNA or can be activated temporarily. Such temporary activation is useful in the production of the crystalline toxin of *Bacillus thuringiensis*, the production of herbicide resistance or induction of resistance to pathogens.

Detailed Description of the Invention

Chromosomal DNA isolated from purified nuclei of soybeans was subjected to restriction enzyme analysis and southern blot hybridizations using cloned hs-cDNAs as hybridization probes. Six sets of cDNA clones which reflected members of "different" multigene families for hs proteins ranging from 15 to 27 kD (there are at least 30 hs proteins in plants within this size range) have been identified. In addition, one gene clone (hs6871) was isolated by use of cDNA clone 1968 as a probe. These groups are as follows:

Table 1

| Classification of heat shock gene families in soybeans ^{***} | | | |
|---|--------------------------------------|-------------------------|------------------------------|
| cDNA clone | Number of Clones by Hybrid Selection | Translation of hs mRNAs | Heat shock promoter fragment |
| pFS2005* | 13 | 15-18 kD hs proteins | hsp2005 |
| pFS2019 ^{***} | 1 | 18 kD hs protein | hsp2019 |
| pCE53* | 13 | 15-18 kD hs proteins | hsp53 |
| pCE75 | 8 | 15 kD hs proteins | hsp75 |
| pCE54 ^{***} | 4(5) | 27 kD hs proteins | hsp54 |
| pFS2033 | 3 | 20-24 kD hs proteins | hsp2033 |
| hs6871 (gene clone) | 1 | 173 kD hs protein | hsp6871 |

* These and several other cDNA clones reflect members of the same multigene family.

^{***} May reflect also a "unique" cDNA clone to the same multigene family represented by pFS2005 and pCE53.

^{***} The expression of these heat shock gene families is induced following many other stress treatments, e.g., treatment with arsenite.

A major effort has been devoted to isolating, restriction mapping, subcloning and sequencing genomic clones isolated from a λ 1059 soybean library (constructed by the Agrigenetics Advanced Research Laboratory, Madison, Wisconsin), and to DNA sequence analysis of the cDNA clones described above.

Four subclones, namely pE2019, pL2005, pM2005 and hs6871 were generated from λ genomic clones and were sequenced. These four subclones represent three different gene families, i.e., cDNA clones 2005, 2019 and 1968. EcoRI fragments were selected on the basis of their homology to cDNAs from these genomic clones and were subcloned into pUC9. Sequence analysis was done on parts of these inserts, i.e., 858 nucleotides of pE2019 (Fig. 1), 971 nucleotides of pL2005, 944 nucleotides of pM2005, and 1536 nucleotides of hs6871, respectively. Sequence homology analysis has located the corresponding cDNA homology within each of the genomic subclone sequences. Cross-homology analysis showed substantial cDNA homology among the genomic subclones: (a) pE2019 corresponds to greater than 98% homology over the 340 bases of pFS2019 cDNA clone (b) immediately upstream from this 2019 homology is a region which shows over 90% homology to the pCE53 cDNA clone sequence of 350 bases. Therefore, sequence data and hybrid selection/translation data indicate that pCE53 is a member of the pFS2005 gene family (see Table 1 above) (c) Additionally, pFS2019 was found to be a 3'-nontranslated region of a member of the 2005 gene family. (d) Homology matrix analysis has shown that even cDNAs that do not cross hybridize (e.g., pCE75 and pFS2033) contained long stretches of greater than 50% homology with 70% homology over lengths of up to 40 nucleotides.

A 7 kilobase HindIII fragment (H2) (Fig. 2) of the soybean genomic clone pE2019 that hybridizes to the cDNA 2019 has been identified. The pE2019 gene was mapped to the left terminal BamHI subfragment (designated BH - Fig. 2) of H2 since this region was shown to be the only portion hybridizing to the cDNA 2019 probe. This fragment was selected for sequencing studies and for introduction into the Ti-plasmid. Using the M13 kilodeletion strategy (Barnes, W. M. and M. Bevan (1983) Nucleic Acids Research 11:349-368) fragment BH has been sequenced and a series of deletion phages spanning this subfragment have been constructed. Of interest are several regions showing 78% homology with the Drosophila heat shock consensus sequence (CTxGAAxxTTCxAG) (Fig. 3) and the core (73%) of the SV40 base pair repeat (GGTGTGGAAAG) (Fig. 3).

The polarity of the transcript and the position of the cDNA homology on H2 (see Fig. 4 for coding sequence) demonstrated that the promoter to gene 2019 is situated on H2 about one hundred and ninety base pairs to the right of the BamHI site of subfragment BH. The polarity of the 2019 gene is 5'- to 3'- from one hundred and fifty seven nucleotides upstream from the leftmost BamHI site of H2 toward the left HindIII site. This conclusion is based on the hybridization of M13 single stranded probes and S1 hybrid protection

studies with soybean heat shock RNA. The 5'-termini of all three genes are positioned from 32 to 28 bases from the first T of a "TATA" motif (the TTAAATAC) suggesting that this region functions as the promoter. Northern blot analysis and S1 hybrid protection results using heat shock RNA showed that the transcript was from 680-800 bases in length excluding the poly(A) tail which was shown to be approximately 150 bases in length. Since the cDNA was less than full length (Ca. 350 bases) and was obtained by priming with oligo-dT, it therefore represents the 3'-portion of the transcript. Therefore, from the position of the cDNA homology on the BH fragment, the transcript must extend 3'- to 5'- from position 590 towards the BamHI site and beyond. This conclusion was confirmed by S1 hybrid protection mapping using 3'-labelled BH fragment. A protected band of 590 ± 10 base pairs was seen which agreed with the 3'-end of cDNA homology and showed that the 5'-terminus and promoter lay to the right of the BamHI site.

The coding sequence of clone pE2019 has been completed (Fig. 4). It consists of an open reading frame of 462 nucleotides. In addition, 291 nucleotides on the 5'-side (i.e., upstream) of the ATG translation initiation codon have been sequenced. These 291 nucleotides include all the essential elements of a promoter region, i.e., CAAT box, TAATA box and transcription initiation. In addition, there is a "consensus sequence" (131-144 nucleotides upstream from the ATG translation initiation codon with the sequence 5'-CTxGAAxxTTCxAG-3') which has been found in all heat shock genes and is required for heat induction (Pelham, H. R. B. and M. Bienz (1982) EMBO J. 1:1473-1477). If this consensus sequence has been deleted from the promoter region, then the heat shock gene is not induced following the stress of a heat shock or any other stress. Another sequence which has a high homology to the SV-40 enhancer sequence occurs at 172 to 185 nucleotides upstream from the translation initiation codon but, at present, the significance of this discovery is not obvious. A conserved sequence is present far upstream and this sequence is also found in an analogous position in two *Drosophila* heat shock promoters. Finally, the sequence of 731 nucleotides on the 3'-side (i.e., downstream) from the TGA stop codon has been elucidated (part of this sequence is shown in Fig. 5).

The coding sequences and flanking sequences of three other heat shock genes (i.e., clones pM2005, pL2005 and hs8871) have been determined. In the case of pM2005, an open reading frame of 423 nucleotides (Fig. 4) has been determined as well as 418 nucleotides upstream (and including all the promoter regulatory sequences described for pE2019 in the previous paragraph) from the ATG translation initiation codon (Fig. 3) and 171 nucleotides downstream from the TAA stop codon (100 nucleotides shown in Fig. 5). In the case of pL2005, an open reading frame of 450 nucleotides (Fig. 4) has been determined as well as 422 upstream (Fig. 3) (and including all the promoter regulatory sequences described for pE2019 in the previous paragraph) from the ATG translation initiation codon and 842 nucleotides downstream from the TAA stop codon (100 nucleotides shown in Fig. 5). In the case of hs8871, an open reading frame of 459 nucleotides (Fig. 4) has been determined as well as 456 upstream (Fig. 3) (and including all the promoter regulatory sequences described for pE2019 in the previous paragraph) from the ATG translation initiation codon and 943 nucleotides downstream from the TAA stop codon (100 nucleotides shown in Fig. 5).

These four heat shock genes have substantially homologous sequences in the coding regions (Fig. 4). In the upstream promoter regions (Fig. 3) the clones pE2019, pM2005 and pL2005 have substantially homologous sequences, but there are many differences between the nucleotide sequences of these three clones and that of hs8871. However, it should be noted that there are strong similarities between the "heat shock consensus sequences" of all four clones, i.e., CTxGAAxxTACxxx (Fig. 3). The data for 100 nucleotides on the downstream of the stop codons for the four sequences are given (Fig. 5). It is obvious that very little sequence homology occurs. Significantly, the coding sequences, the upstream promoter regions (i.e., 5'- to the translational initiation codon) and the downstream flanking region (i.e., 3'- to the stop codon) of these four soybean heat shock genes have almost no resemblance to the corresponding regions of *Drosophila* heat shock genes (Hackett, R. W. and J. T. Lis (1983) Nucleic Acids Res. 11:7011-7030; Ingolia, T. D. and E. A. Craig (1982) Proc. Nat. Acad. Sci. USA 79:2360-2364; Southgate, R. et al. (1983) J. Mol. Biol. 165:35-67). Although there are similarities between the "consensus sequences" of the promoter regions from *Drosophila* and soybean heat shock genes, the promoter regions of soybean heat shock genes do not possess the inverted repeat sequences characteristic of the *Drosophila* genes.

The promoter regions of the soybean heat shock genes can be utilized in a number of ways whenever a transitory activation of a foreign gene or a soybean gene is required. [A foreign gene is herein defined as any gene normally found in the genome of any species other than a soybean.] For example, when the T-DNA from a wild type Ti-plasmid of *Agrobacterium tumefaciens* is transferred to a plant genome, then the resultant transformed plant cells are tumorous. These transformed tumorous plant cells in tissue culture cannot be used to regenerate intact whole plants. On the other hand, if a "disarmed" T-DNA region is used, then intact whole plants can be regenerated from transformed plant cells in tissue culture, but it is difficult to differentiate between transformed and untransformed cells. In the present invention this difficulty is

overcome by placing the β -galactosidase gene of *E. coli* under the heat shock inducible control of a soybean heat shock gene promoter. This recombinant construction comprises the soybean heat shock promoter region and the coding region for the 24 codons at the 5'-terminus of the heat shock gene (see Example 5). This recombinant DNA fragment is then integrated into the T-DNA of a Ti-plasmid and used to transform plant cells. In the presence of a suitable substrate, transformed cells in tissue culture can then be differentiated from untransformed cells by the development of a blue color following a heat shock treatment. Thus the β -galactosidase-heat shock promoter combination is used as a means of recognizing transformed plant cells. This invention is not limited to the β -galactosidase gene example and will include other genes which could be useful in the recognition of specific plant cell types when such genes are placed under the control of a plant heat shock gene promoter. Such genes useful in the recognition of transformed plant cells are defined herein as transformation recognition genes.

In a second example, it will be useful to follow the protocol outlined above except that a gene which one wishes to be transiently expressed is placed under the control of the soybean heat shock promoter. This recombinant construction comprises only the soybean heat promoter region, i.e., 159 nucleotide pairs extending from the *AluI* site (17 base pairs upstream from the ATG translation initiation codon) to the *EcoRI* site (176 base pairs upstream from the ATG translational initiation site) (see Example 7). If genes coding for insecticidal proteins (including but not limited to the crystalline endotoxin of *Bacillus thuringiensis*) are placed under the control of a plant heat shock gene promoter, then it will be possible to activate expression of the insecticidal proteins during the heat of the day. This period coincides with the eating period of insect larvae and thus confers insect resistance to the plant—but only for a limited, critical period during each day. Similarly, if a gene conferring herbicide resistance is placed under the control of a heat shock promoter, then it is possible to spray the fields with herbicide after the herbicide resistance gene has been activated during the heat of the day.

25 Example 1: Plant material used and heat shock conditions

Soybean seeds (*Glycine max* variety Wayne) were germinated in moist vermiculite in the dark at 28°-30° C for 3 days. After this time plants were sprayed with 2×10^{-3} M 2,4-dichlorophenoxyacetic acid, and mature hypocotyl tissue was harvested 24 hours later. The tissue was incubated in a buffer containing 1% sucrose, 1 mM K-phosphate (pH 6.0), 50 μ g/ml chloramphenicol 10 μ g/ml 1,4-dichlorophenoxyacetic acid at 28° C (control) or at 40° C and 42.5° C (heat shock), respectively, for two hours unless stated differently.

Example 2: Purification of poly(A)⁺ RNA and construction of cDNA recombinant clones

Total RNA was extracted from hypocotyl tissue after incubation (see Example 1), and poly(A)⁺ RNA was purified as described (Silflow, C. D. et al. (1979) *Biochemistry* 13:2725-2731) with modifications (Key, J. L. et al. (1981) *Proc. Nat. Acad. Sci. USA* 78:3526-3530). Poly(A)⁺ RNA from soybean hypocotyl was used as a template for oligo-(dT)-primed double stranded cDNA synthesis (Wickens, M. P. et al. (1978) *J. Biol. Chem.* 253:2483-2495; modified by Baulcombe, D. C. and J. L. Key (1980) *J. Biol. Chem.* 255:8907-8913). As a further modification, the synthesis of the first strand was unlabelled and 20 μ M [³²P]dCTP (400 Ci/mM, Amersham) was used as a tracer for second strand synthesis by DNA polymerase I (Boehringer Mannheim). The S1-digested double stranded cDNA was size fractionated on a 10-30% sucrose gradient in 10 mM Tris-HCl pH 7.5, 1 mM EDTA and 100 mM NaCl, run at 50,000 rpm for 6 hours at 20° C in a Beckman SW 50.1 rotor. About 0.5 μ g double stranded cDNA (of length greater than 500 bp) was subjected to homopolymer tailing, adding poly(dC) to the 3'-ends of fragments by terminal transferase (Bethesda Research Laboratories) (Roychoudhury, R. and R. Wu (1980) In: Grossman, L., Moldave, K., eds. *Methods in Enzymol.* Vol. 65; New York: Academic Press, pp. 43-62). An average length of 30 nucleotides/end was synthesized. In an analogous reaction 1 μ g of *Pst*I cut pBR322 was tailed with poly(dG) to the same extent. In an annealing reaction 0.7 μ g (dG)-tailed pBR322 and 0.14 μ g (dC)-tailed cDNA were used. Annealed molecules were used to transform *Escherichia coli* SK1590 (Kushner, S. R. (1978) In: Boyer, H. W., Nicosia, S. eds. *Genetic Engineering*, Amsterdam: Elsevier/North Holland Biomedical Press, pp. 17-23). Transformants were selected on tetracycline-containing medium, 99% of which carried recombinant plasmids as indicated by their TcR ApS phenotype.

55 Example 3: Screening of a soybean genomic DNA library

High molecular weight DNA was isolated from purified nuclei essentially as described (Nagao, R. et al. (1981) *DNA* 2:1-9). Screening of a soybean genomic DNA library, cloned into the *EcoRI* site of a λ Charon

4A vector was carried out as described (Nagao, R. T. et al. (1981) DNA 1:1-9), using radioactively labelled insert probes of cDNA clones which had been synthesized from poly(A)⁺ RNA of heat shock treated soybean hypocotyls (Schoffl, F. and J. Key (1982) J. Mol. Appl. Genet. 1:301-314).

5 Example 4: Restriction endonuclease digestion and subcloning of DNA fragments in pBR322

Assay conditions for DNA digestions with the restriction endonucleases EcoRI, HindIII and PstI were as described (Maniatis, T. et al. (1982) Molecular cloning, a Laboratory Manual, Cold Spring Harbor Laboratory) and standard electrophoresis of DNA fragments on 1% agarose gels was also as described (Schoffl, F. and A. Puhler (1979) Genet. Res. Camb. 34:287-301). Ten µg/lane was applied for soybean chromosomal DNA digests and about 0.5µg/lane for plasmid or λ-DNA digests. Completion of digestion was tested for soybean chromosomal DNA by southern blot hybridization with soybean rDNA probes (kindly provided by Dr. R. Nagao, University of Georgia). Fragment sizes were generally determined by comparison with λ-DNA digests (EcoRI, HindIII, EcoRI/HindIII) run on the same gel. Subcloning of EcoRI/HindIII fragments of genomic soybean DNA into the respective sites of pBR322 was carried out as described (Maniatis, T. et al. (1982) supra). Potential recombinant clones were screened by sizing the cloned DNA fragments on agarose gels using restriction fragments of a standard heat shock gene (λhs68-7) as a reference. Specific clones were identified by Southern blot hybridization using cDNA probes of clone 1968 (Schoffl, F. and J. L. Key (1982) J. Mol. Appl. Genet. 1:301-314).

20 Example 5: Construction of a recombinant plasmid containing the β-galactosidase gene inserted into the coding region of soybean heat shock gene 2019

The starting material for this construction (defined here as a recombinant soybean heat shock gene) is the 7 kilobase (kb) HindIII fragment (H2) containing the promoter of heat shock gene 2019, the coding sequence of the same heat shock gene and a flanking sequence on the 3'-side of the reading strand of this coding sequence (referred to hereafter as hs2019) (Fig. 6). This H2 sequence is digested with the restriction endonuclease EcoRI and the products are separated by electrophoresis on agarose gel. The 1.78 kb HindIII-EcoRI which contains all the components of heat-shock gene 2019 is then inserted into plasmid pBR322 previously cleaved with HindIII and EcoRI. This recombinant plasmid (pBR322-hs2019) is then transformed into E. coli JM101 where it is amplified. Following amplification, pBR322-hs2019 is partially cleaved with BamHI and a Z gene (coding for β-galactosidase) carrying polylinkers at both ends (Casadaban, M. J. et al. (1983) Methods Enzymol. 100:293-308) is inserted into the hs2019 BamHI site. The polylinkers on each end of the Z-gene are previously cleaved with BamHI. It should be noted that the Z-gene may also insert into the BamHI site of pBR322. Insertion at the two sites can be differentiated by restriction mapping. The BamHI site of the hs2019 gene is at codon 24 of the hs2019 coding region, and this construction thus maintains the coding region of the Z-gene inframe following the first 24 codons of the hs2019 gene. This recombinant soybean heat shock gene inserted into pBR322 is designated hereinafter as pBR322/SB13 (i.e., pBR322 with an hs2019 promoter-24 codons of hs 2019 coding sequence-Z-gene-hs2019 coding sequence-3'-flanking sequence). SB13 (referred to hereinafter as a recombinant soybean heat shock gene) can be recovered from pBR322/SB13 by cleavage with HindIII and EcoRI followed by separation of the products by agarose gel electrophoresis.

45 Example 6: Incorporation of soybean hs-promoter-β-galactosidase-hs coding-hs3'-tail (i.e., SB13) into the Ti-plasmid of Agrobacterium tumefaciens

A T-DNA shuttle vector p233G comprising pBR322 and the T_L-DNA of the Ti-plasmid of A. tumefaciens was obtained from the Agrigenetics Advanced Research Laboratory, Madison, Wisconsin. This T-DNA shuttle vector (p233G) had been transformed into E. coli JM101. pBR322 is resistant to both ampicillin (amp^r) and tetracycline (tet^r), but p233G is only amp^r because the T_L-DNA has been inserted into the tet^r gene, thus destroying its activity (Fig. 7). Following amplification, p233G is purified and cleaved at the SmaI restriction endonuclease site in transcript number 10 of T-DNA from A. tumefaciens strain 15955 (Barker, R. F. et al (1983) Plant Mol. Biol. 2:335-350). BglII linkers are then added to this SmaI site and digested with BglII.

SB13, previously recovered by cleavage of pBR322/SB13 by HindIII and EcoRI, has overhanging single stranded ends produced by the action of these restriction endonucleases. These overlaps are filled in by use of DNA polymerase I (Klenow fragment) and BglII linkers are blunt end ligated (Fig. 7). SB13 is then digested with BglII restriction endonuclease and the BglII fragment is isolated by agarose gel elec-

trophoresis. It will be noted that a BglII restriction site occurs 39 base pairs from the 3'-terminus of the hs2019 mRNA (Fig. 2) and that this BglII site will therefore represent one end of the BglII fragment. For this reason, the nomenclature of the SB13 fragment is altered to SB13'. The BglII fragment is then inserted into the linearized p233G, i.e., the T-DNA shuttle vector, resulting in plasmid p233G/SB13' (defined here as a co-integrated recombinant DNA fragment) and transformed into *E. coli* strain JM101. Following amplification, a triple mating is done using (1) a helper plasmid (pRK2013) in an *E. coli* strain, (2) the plasmid p233G/SB13' in *E. coli* JM101 and (3) strain 15955 of *A. tumefaciens* containing a Ti-plasmid (Fig. 8). Strain 15955 is resistant to streptomycin (str^r). pRK2013 and the recombinant shuttle vector p233G/SB13' have replication origins which are functional in *E. coli* strains but not in *A. tumefaciens*. Thus a helper plasmid can be defined as a plasmid that promotes transfer of a normally non-transferable second plasmid from one bacterial strain to another. However, pBR322 has a mobilization site (mob) which is recognized by the transfer gene (tra) of pRK2013 so the recombinant shuttle vector p233G/SB13' can be transferred to *A. tumefaciens*. However, p233G/SB13' cannot replicate in *A. tumefaciens*, so its presence can only be stabilized by recombination (single crossover or double reciprocal crossover) with the resident Ti-plasmid. These three strains are mixed and incubated for 16 hours after which the recombinant resident Ti-plasmid (i.e., Ti-p233G/SB13') is selected by plating for 72 hours on a medium containing streptomycin and carbenicillin. Streptomycin selects the *Agrobacterium* and carbenicillin selects the pBR322. The recombinant Ti-p233G/SB13' promoter plasmid in *A. tumefaciens* strain 15955 can now be utilized.

The recombinant Ti-p233G/SB13' resident plasmid now contains the β -galactosidase producing gene (i.e., the Z-gene) under the control of the hs2019 heat shock promoter within the T-DNA of the Ti-plasmid in a stable form within *A. tumefaciens* strain 15955. Following infection of a plant or plant cell culture by the bacteria, the T-DNA can be transferred to the plant genome. Plant tissue or plant cells, which have been thus transferred, can then be recognized by the expression of the Z-gene (defined here as a transformation recognition gene) resulting in the production of a blue color after heat treatment in a medium containing 5-bromo-4-chloro-3-indolyl- β -D-galactoside (X-gal) (Miller, J. H. (ed.) (1975) Experiments in Molecular Genetics. Cold Spring Harb. Lab., Cold Spring Harbor, New York). Most important is that the expression of the blue color is only transitory.

Example 7: Isolation of the heat shock promoter from heat shock gene 2019 of soybeans and insertion of this heat shock promoter into T₁ DNA of plasmid p233G

Starting material for the isolation of gene 2019 heat shock promoter is the pUC8-derived clone BE250 (Fig. 2). The plasmid pUC8-BE250 contains the BamHI-EcoRI subfragment of H2 that includes the promoter and part of the coding region of the heat shock gene 2019. This plasmid is digested with restriction endonuclease AluI and the promoter-containing fragment is isolated (Fig. 1). The fragment extends 65 base pairs downstream from the start of transcription to include a major portion of the untranslated leader sequence, but not the start codon for translation. HindIII linkers are blunt-end ligated to the fragment and the ligation product is redigested with HindIII and BamHI before cloning into similarly digested pUC8. Heat shock promoters from all three isolated genes (E2019, M2005, L2005) are cloned in a similar fashion and designated hspE2019, hspM2005, and hspL2005, respectively.

The plasmid p233G is linearized by digestion with BglII and the resulting single-strand end filled in using the Klenow fragment of DNA polymerase I. A synthetic polylinker (5'-GAGATCTAAGCTTCTAGAC-3', double stranded) is ligated into p233G at the filled-in BglII site. This polylinker contains the restriction sites of BglII, HindIII and XbaI endonucleases, and is used for insertion of both the BamHI/HindIII-flanked promoter fragments and HindIII/XbaI-generated coding region fragments. The coding region fragments can be obtained from any gene as long as the fragments contain no upstream ATG sequences other than the start codon for translation. The coding fragment must also contain an untranslated 3'-tail with a polyA addition site (AATAAA) for correct processing of the mRNA.

Such heat shock expression plasmids are then transformed into a strain of *E. coli*, e.g., JM101 or JM103, which will allow replication. Following amplification in such a host strain, the heat shock expression plasmids can be transferred to *Agrobacterium* strains which can then be used to transform plant cells as already described in Example 6.

Example 8: Construction and isolation of a heat shock promoter sequence from the heat shock gene clone phs6871

The isolation of a soybean heat shock gene and the insertion of this gene into pBR322 to give a recombinant plasmid phs6871 have been described (Schoffl, F. and J. L. Key (1983) Plant Mol. Biol. 2:269-

278). Following amplification in an *E. coli* K12 strain, the recombinant plasmid is purified and cleaved with a mixture of the restriction endonucleases *Cfr*I and *Acc*I. If the first nucleotide A of the ATG translation initiation codon is numbered +1, then cleavage with these two restriction endonucleases will produce a fragment covering the nucleotides -314 (i.e., 314 nucleotides "upstream" from the above-described A) (Fig. 9). Upstream is defined here as being in the 5'-direction and downstream is in the 3'-direction from the A nucleotide of the ATG translation initiation codon on the reading strand of the DNA. The fragment is then purified and the single stranded overhangs created by the restriction enzymes are blunt-ended by methods well known in the art (Maniatis T. et al. (1982) Molecular cloning - a laboratory manual. Cold Spring Harbor Laboratory). *Eco*RI linkers are then added to both ends of the fragment which is then cloned into the *Eco*RI site of M13mp9. Following transformation into *E. coli* JM103, the cloned fragment is amplified and single stranded templates corresponding to the reading strand of the heat shock gene are packaged and extruded into the media. These single stranded templates are recovered from the supernatant following removal of the bacterial host. A ten-fold excess of a previously constructed synthetic DNA primer mismatched in four base pairs (5'-TTTCCCGGGTCAGTCTTG-3') in the presence of the four deoxynucleotide triphosphates (one of which is radioactive) and DNA polymerase I (Klenow fragment) is now used to generate a modified double stranded DNA.

The four mismatched nucleotides CCGGG are indicated by underlines. The mixture is incubated for a sufficient period at 37°C to allow two full cycles of replication. The fragment containing the *hs6871* promoter region is then isolated and purified following a mixed digestion with the restriction endonucleases *Eco*RI and *Sma*I. The overhang generated by the *Eco*RI digestion is then blunt ended and the fragment (now with both ends blunt) is then blunt end ligated into the *Sma*I site of p233G and the recombinant DNA plasmids are amplified following transformation into a suitable host. It will be noted that two full cycles of DNA replication starting from the mismatched synthetic DNA primer generates a *Sma*I restriction site. The fragment containing the *hs6871* promoter will be inserted in both orientations into the *Sma*I site of p233G, but in both orientations a *Sma*I site is regenerated downstream from the *hs6871* promoter sequence and can be utilized for the insertion of foreign genes or soybean genes of interest. In particular, it should be noted that there is NO *Sma*I site generated upstream from the *hs6871* promoter region. As described in Example 7, the p233G constructs containing an inserted heat shock promoter *hs6871* are defined as recombinant DNA plasmids. These recombinant DNA plasmids with foreign genes or soybean genes inserted are designated (as in Example 7) as heat shock expression plasmids. Transfer of heat shock expression plasmids into a plant genome is accomplished as described in Example 7.

Claims

1. A recombinant DNA plasmid comprising:
 - (a) a vector,
 - (b) a T-DNA fragment of a Ti-plasmid from an *Agrobacterium* strain, which T-DNA fragment is functional for insertion into a plant genome, and
 - (c) a fragment of plant DNA which is capable of controlling gene expression in response to heat shock, and which fragment is inserted into said T-DNA fragment such that said T-DNA fragment retains functionality for insertion into a plant genome, wherein said fragment of plant DNA comprises a plant heat shock promoter comprising the consensus nucleotide sequence 5'-C-T-X-G-A-A-X-X-T-A-C-X-X-X-3', wherein X is A, T, C or G.
2. A recombinant DNA plasmid according to claim 1 wherein said promoter comprises the consensus nucleotide sequence 5'-C-T-S-G-A-A-M-R-T-A-C-W-M-K-3', where S is C or G; M is A or C; R is A or G; W is A or T; and K is T or G.
3. A recombinant DNA plasmid according to claim 1 wherein said heat shock promoter is derived from soybean.
4. A recombinant DNA plasmid according to claim 3 wherein said heat shock promoter is a soybean heat shock promoter selected from the group consisting of *hsp2005*, *hsp2019* and *hsp6871*.
5. A recombinant DNA plasmid according to claim 1 wherein said fragment of plant DNA is obtained from a soybean plant, the fragment thereof comprising the sequence

5'-T-A-C-A-T-

G-G-T-G-T-G-G-A-G-A-A-T-T-C-A-A-C-C-A-A-A-T-T-G-C-A-A-A-A-
 A-G-T-A-G-G-A-T-T-T-T-T-C-T-G-G-A-A-C-A-T-A-C-A-A-G-A-T-T-
 5 A-T-C-C-T-T-T-C-A-C-T-T-C-C-T-T-T-A-A-A-T-A-C-C-T-C-G-C-G-
 T-A-T-C-C-C-C-T-T-C-G-T-C-C-T-C-G-T-C-A-A-A-C-G-A-A-G-A-A-
 A-A-A-A-G-T-T-A-C-C-T-G-T-T-T-G-C-G-A-T-C-T-C-A-T-T-A-C-A-
 10 A-T-C-T-C-C-C-T-A-G-T-T-T-C-T-A-A-T-C-T-C-A-G-3',

or being substantially homologous therewith.

- 15 6. A recombinant DNA plasmid according to claim 1 wherein said fragment of plant DNA is obtained from a soybean plant, the fragment thereof comprising the sequence

5'-T-A-C-A-T-

G-G-T-G-T-G-G-A-G-A-A-T-T-C-A-A-C-C-A-A-A-T-T-G-C-A-A-A-A-
 A-G-T-A-G-G-A-T-T-T-T-T-C-T-G-G-A-A-C-A-T-A-C-A-A-G-A-T-T-
 A-T-C-C-T-T-T-C-A-C-T-T-C-C-T-T-T-A-A-A-T-A-C-C-T-C-G-C-G-
 20 T-A-T-C-C-C-C-T-T-C-G-T-C-C-T-C-G-T-C-A-A-A-C-G-A-A-G-A-A-
 25 A-A-A-A-G-T-T-A-C-C-T-G-T-T-T-G-C-G-A-T-C-T-C-A-T-T-A-C-A-
 A-T-C-T-C-C-C-T-A-G-T-T-T-C-T-A-A-T-C-T-C-A-G-C-T-A-A-G-A-
 A-A-A-A-C-C-A-A-A-A-G-A-T-G-T-C-T-C-T-G-A-T-T-C-C-A-G-G-T-
 30 T-C-C-T-T-C-G-G-T-G-G-C-C-G-A-A-G-G-A-G-C-A-A-C-G-T-C-T-T-
 C-G-A-T-C-C-A-T-T-C-T-C-A-C-T-C-G-A-C-A-T-G-T-G-G-3',

35 or being substantially homologous therewith.

7. A recombinant DNA plasmid according to claim 1 wherein said fragment of plant DNA is obtained from a soybean plant, the fragment thereof comprising the sequence

5'-A-G-A-C-C-

A-A-T-C-C-T-A-A-C-C-A-A-T-G-T-C-T-G-G-T-T-A-A-G-A-T-G-G-T-
 C-C-A-A-T-C-C-C-G-A-A-A-C-T-T-C-T-A-G-T-T-G-C-G-G-T-T-C-G-
 45 A-A-G-A-A-G-C-C-A-G-A-A-T-G-T-T-T-C-T-G-A-A-A-G-T-T-T-C-A-
 G-A-A-A-A-T-T-C-T-A-G-T-T-T-T-G-A-G-A-T-T-T-T-C-A-G-A-A-G-
 T-A-C-G-G-C-A-T-G-A-T-G-A-T-G-C-A-T-A-A-C-A-A-G-G-A-C-T-T-
 50 T-C-T-C-G-A-A-A-G-T-A-C-T-A-T-A-T-T-G-C-T-C-C-T-C-T-A-C-A-
 T-C-A-T-T-T-T-A-A-A-T-A-C-C-C-C-A-T-G-T-G-T-C-C-T-T-T-G-A-
 A-G-A-C-A-C-A-T-C-A-C-A-G-A-A-A-G-A-A-G-T-G-A-A-G-G-C-A-T-
 55 C-G-T-T-A-G-C-A-G-T-T-T-T-G-T-A-G-A-T-T-C-A-A-C-C-T-C-A-A-
 T-T-T-G-C-A-G-A-G-T-T-A-C-G-T-T-C-T-A-A-T-A-T-T-T-T-A-C-
 A-C-A-A-G-A-C-T-G-A-C-C-C-3',

or being substantially homologous therewith.

8. A recombinant DNA plasmid according to claim 1 wherein said fragment of plant DNA is obtained from a soybean plant, the fragment thereof comprising the sequence

5

5'-A-G-A-C-C-

A-A-T-C-C-T-A-A-C-C-A-A-T-G-T-C-T-G-G-T-T-A-A-G-A-T-G-G-T-
 C-C-A-A-T-C-C-C-G-A-A-A-C-T-T-C-T-A-G-T-T-G-C-G-G-T-T-C-G-
 10 A-A-G-A-A-G-C-C-A-G-A-A-T-G-T-T-T-C-T-G-A-A-A-G-T-T-T-C-A-
 G-A-A-A-A-T-T-C-T-A-G-T-T-T-T-G-A-G-A-T-T-T-T-C-A-G-A-A-G-
 T-A-C-G-G-C-A-T-G-A-T-G-A-T-G-C-A-T-A-A-C-A-A-G-G-A-C-T-T-
 15 T-C-T-C-G-A-A-A-G-T-A-C-T-A-T-A-T-T-G-C-T-C-C-T-C-T-A-C-A-

 T-C-A-T-T-T-T-A-A-A-T-A-C-C-C-C-A-T-G-T-G-T-C-C-T-T-T-G-A-
 A-G-A-C-A-C-A-T-C-A-C-A-G-A-A-A-G-A-A-G-T-G-A-A-G-G-C-A-T-
 C-G-T-T-A-G-C-A-G-T-T-T-T-G-T-A-G-A-T-T-C-A-A-C-C-T-C-A-A-
 T-T-T-G-C-A-G-A-G-T-T-A-C-G-T-T-C-T-A-A-T-A-T-A-T-T-T-A-C-
 25 A-C-A-A-G-A-C-T-G-A-T-A-A-G-A-G-A-A-A-A-T-G-T-C-T-C-T-G-A-
 T-T-C-C-A-A-G-T-T-T-C-T-T-C-G-G-T-G-G-C-C-G-A-A-G-G-A-G-C-
 A-G-T-G-T-T-T-T-C-G-A-C-C-C-T-T-T-C-T-C-C-C-T-C-G-A-T-G-T-
 30 G-T-G-G-3',

or being substantially homologous therewith.

- 35 9. A recombinant DNA plasmid according to any preceding claim wherein said vector is pBR 322.
10. A recombinant DNA plasmid according to any preceding claim further comprising a structural gene under the control of said fragment of plant DNA, which structural gene is not a heat shock gene naturally found with said heat shock promoter.
- 40 11. A recombinant DNA plasmid according to claim 10 wherein said structural gene is a foreign gene.
12. A recombinant DNA plasmid according to claim 11 wherein said foreign structural gene is a herbicide resistance gene.
- 45 13. A recombinant DNA plasmid according to claim 11 wherein said foreign gene is a Z-gene coding for β -galactosidase.
14. A recombinant DNA plasmid according to claim 11 wherein said foreign gene is a crystalline endotoxin gene.
- 50 15. A recombinant DNA plasmid according to claim 14 wherein said crystalline endotoxin gene is a crystalline endotoxin gene of Bacillus thuringiensis.
- 55 16. A recombinant DNA plasmid according to claim 10 wherein said structural gene is a soybean gene.
17. A method for modifying a plant genome so as to allow expression of a structural gene under control of a soybean heat shock gene promoter fragment, comprising the steps of:

- (a) isolating a soybean heat shock gene promoter fragment comprising the consensus nucleotide sequence 5'-C-T-X-G-A-A-X-X-T-A-C-X-X-X-3', wherein X is A, T, C or G,
 - (b) cloning said soybean heat shock gene promoter fragment into a T-DNA shuttle vector producing a recombinant DNA plasmid,
 - 5 (c) isolating a DNA fragment carrying foreign structural genes or soybean genes not naturally expressed under the regulatory control of a soybean heat shock gene promoter, and inserting said DNA fragment into said recombinant DNA plasmid at a position on the 3'-side of said soybean heat shock gene promoter producing a heat shock expression plasmid, wherein said DNA fragment is oriented with respect to said soybean heat shock gene promoter as to be expressible under control
 - 10 thereof,
 - (d) transforming said heat shock expression plasmid into a first bacterial strain capable of supporting replication of said heat shock expression plasmid,
 - (e) mixing said bacterial strain capable of supporting replication of said heat shock expression plasmid with a second bacterial strain carrying a helper plasmid capable of transporting said heat
 - 15 shock expression plasmid into an Agrobacterium strain incapable of supporting replication of said heat shock expression plasmid, said Agrobacterium strain carrying a resident plasmid,
 - (f) selecting for recombination between said heat shock expression plasmid and said resident plasmid giving a recombinant resident plasmid,
 - (g) infecting a plant or a plant cell culture with said Agrobacterium strain containing and replicating
 - 20 said recombinant resident plasmid, and
 - (h) selecting a plant or a plant cell culture comprising plant cells containing said foreign structural genes or said soybean genes under control of said soybean heat shock gene promoter transferred from said recombinant resident plasmid to said plant cells, said foreign structural genes or said soybean genes being expressed following heat shock treatment or other stress treatment.
 - 25
18. A method according to claim 17 wherein said heat shock gene promoter fragment is as recited in any of claims 2 to 8.
19. A method according to claim 17 or claim 18 wherein the DNA fragment isolated in step (c) comprises a
- 30 foreign gene according to any of claims 12 to 15.
20. A method for producing a genetically modified plant cell, said method comprising
- i) infecting a plant or a plant cell culture with an Agrobacterium strain containing and replicating a recombinant DNA plasmid according to any of claims 10 to 16, and
 - 35 ii) propagating plant cells containing said structural gene under control of said fragment of plant DNA.

Patentansprüche

- 40 1. Rekombinantes DNA-Plasmid, enthaltend
 - (a) einen Vektor,
 - (b) ein T-DNA-Fragment eines Ti-Plasmids aus einem Agrobacterium-Stamm, welches T-DNA-Fragment für die Insertion in ein Pflanzengenom funktionell ist, und
 - 45 - (c) ein Fragment einer Pflanzen-DNA, das imstande ist, die Genexpression als Reaktion auf einen Wärmeschock zu steuern, und das in dieses T-DNA-Fragment derart eingesetzt ist, daß das T-DNA-Fragment die Funktionalität zum Einsetzen in ein Pflanzengenom beibehält, wobei dieses Fragment der Pflanzen-DNA einen pflanzlichen Wärmeschock-Promotor enthält, der die Consensus-Nucleotidsequenz 5'-C-T-X-G-A-A-X-X-T-A-C-X-X-X-3' umfaßt, in welcher X für A, T, C oder G steht.
 - 50
2. Rekombinantes DNA-Plasmid nach Anspruch 1, wobei der genannte Promotor die Consensus-Nucleotidsequenz 5'-C-T-S-G-A-A-M-R-T-A-C-W-M-K-3' umfaßt, in welcher S für C oder G, M für A oder C, R für A oder G, W für A oder T und K für T oder G steht.
- 55 3. Rekombinantes DNA-Plasmid nach Anspruch 1, wobei der genannte Wärmeschockpromotor von der Sojabohne stammt.
4. Rekombinantes DNA-Plasmid nach Anspruch 3, wobei dieser Wärmeschockpromotor ein Sojabohnen-

Wärmeschockpromotor, ausgewählt aus der Gruppe hsp2005, hsp2019 und hsp6871, ist.

5. Rekombinantes DNA-Plasmid nach Anspruch 1, wobei das genannte Fragment der Pflanzen-DNA aus einer Sojabohnenpflanze erhalten wird, deren Fragment die Sequenz

5

5'-T-A-C-A-T-

G-G-T-G-T-G-G-A-G-A-A-T-T-C-A-A-C-C-A-A-A-T-T-G-C-A-A-A-A-
 10 A-G-T-A-G-G-A-T-T-T-T-T-C-T-G-G-A-A-C-A-T-A-C-A-A-G-A-T-T-
 A-T-C-C-T-T-T-C-A-C-T-T-C-C-T-T-T-A-A-A-T-A-C-C-T-C-G-C-G-
 15 T-A-T-C-C-C-C-T-T-C-G-T-C-C-T-C-G-T-C-A-A-A-C-G-A-A-G-A-A-
 A-A-A-A-G-T-T-A-C-C-T-G-T-T-T-G-C-G-A-T-C-T-C-A-T-T-A-C-A-
 A-T-C-T-C-C-C-T-A-G-T-T-T-C-T-A-A-T-C-T-C-A-G-3'

20

umfaßt oder im wesentlichen homolog damit ist.

6. Rekombinantes DNA-Plasmid nach Anspruch 1, wobei das genannte Fragment der Pflanzen-DNA aus einer Sojabohnenpflanze erhalten wird, deren Fragment die Sequenz

25

5'-T-A-C-A-T-

G-G-T-G-T-G-G-A-G-A-A-T-T-C-A-A-C-C-A-A-A-T-T-G-C-A-A-A-A-
 30 A-G-T-A-G-G-A-T-T-T-T-T-C-T-G-G-A-A-C-A-T-A-C-A-A-G-A-T-T-
 A-T-C-C-T-T-T-C-A-C-T-T-C-C-T-T-T-A-A-A-T-A-C-C-T-C-G-C-G-
 T-A-T-C-C-C-C-T-T-C-G-T-C-C-T-C-G-T-C-A-A-A-C-G-A-A-G-A-A-
 35 A-A-A-A-G-T-T-A-C-C-T-G-T-T-T-G-C-G-A-T-C-T-C-A-T-T-A-C-A-
 A-T-C-T-C-C-C-T-A-G-T-T-T-C-T-A-A-T-C-T-C-A-G-C-T-A-A-G-A-
 A-A-A-A-C-C-A-A-A-A-G-A-T-G-T-C-T-C-T-G-A-T-T-C-C-A-G-G-T-
 T-C-C-T-T-C-G-G-T-G-G-C-C-G-A-A-G-G-A-G-C-A-A-C-G-T-C-T-T-
 40 C-G-A-T-C-C-A-T-T-C-T-C-A-C-T-C-G-A-C-A-T-G-T-G-G-3'

umfaßt oder im wesentlichen homolog damit ist.

- 45 7. Rekombinantes DNA-Plasmid nach Anspruch 1, wobei das genannte Fragment der Pflanzen-DNA aus einer Sojabohnenpflanze erhalten wird, deren Fragment die Sequenz

50

55

5'-A-G-A-C-C-
 A-A-T-C-C-T-A-A-C-C-A-A-T-G-T-C-T-G-G-T-T-A-A-G-A-T-G-G-T-
 5 C-C-A-A-T-C-C-C-G-A-A-A-C-T-T-C-T-A-G-T-T-G-C-G-G-T-T-C-G-
 A-A-G-A-A-G-C-C-A-G-A-A-T-G-T-T-T-C-T-G-A-A-A-G-T-T-T-C-A-
 G-A-A-A-A-T-T-C-T-A-G-T-T-T-T-G-A-G-A-T-T-T-T-C-A-G-A-A-G-
 10 T-A-C-G-G-C-A-T-G-A-T-G-A-T-G-C-A-T-A-A-C-A-A-G-G-A-C-T-T-
 T-C-T-C-G-A-A-A-G-T-A-C-T-A-T-A-T-T-G-C-T-C-C-T-C-T-A-C-A-
 T-C-A-T-T-T-T-A-A-A-T-A-C-C-C-C-A-T-G-T-G-T-C-C-T-T-T-G-A-
 A-G-A-C-A-C-A-T-C-A-C-A-G-A-A-A-G-A-A-G-T-G-A-A-G-G-C-A-T-
 15 C-G-T-T-A-G-C-A-G-T-T-T-T-G-T-A-G-A-T-T-C-A-A-C-C-T-C-A-A-
 T-T-T-G-C-A-G-A-G-T-T-A-C-G-T-T-C-T-A-A-T-A-T-A-T-T-T-A-C-
 A-C-A-A-G-A-C-T-G-A-C-C-C-3'

20

umfaßt oder im wesentlichen homolog damit ist.

8. Rekombinantes DNA-Plasmid nach Anspruch 1, wobei das genannte Fragment der Pflanzen-DNA aus einer Sojabohnenpflanze erhalten wird, deren Fragment die Sequenz

25

5'-A-G-A-C-C-

30 A-A-T-C-C-T-A-A-C-C-A-A-T-G-T-C-T-G-G-T-T-A-A-G-A-T-G-G-T-
 C-C-A-A-T-C-C-C-G-A-A-A-C-T-T-C-T-A-G-T-T-G-C-G-G-T-T-C-G-
 A-A-G-A-A-G-C-C-A-G-A-A-T-G-T-T-T-C-T-G-A-A-A-G-T-T-T-C-A-
 35 G-A-A-A-A-T-T-C-T-A-G-T-T-T-T-G-A-G-A-T-T-T-T-C-A-G-A-A-G-
 T-A-C-G-G-C-A-T-G-A-T-G-A-T-G-C-A-T-A-A-C-A-A-G-G-A-C-T-T-
 T-C-T-C-G-A-A-A-G-T-A-C-T-A-T-A-T-T-G-C-T-C-C-T-C-T-A-C-A-
 40 T-C-A-T-T-T-T-A-A-A-T-A-C-C-C-C-A-T-G-T-G-T-C-C-T-T-T-G-A-
 A-G-A-C-A-C-A-T-C-A-C-A-G-A-A-A-G-A-A-G-T-G-A-A-G-G-C-A-T-
 C-G-T-T-A-G-C-A-G-T-T-T-T-G-T-A-G-A-T-T-C-A-A-C-C-T-C-A-A-
 45 T-T-T-G-C-A-G-A-G-T-T-A-C-G-T-T-C-T-A-A-T-A-T-A-T-T-T-A-C-
 A-C-A-A-G-A-C-T-G-A-T-A-A-G-A-G-A-A-A-A-T-G-T-C-T-C-T-G-A-
 T-T-C-C-A-A-G-T-T-T-C-T-T-C-G-G-T-G-G-C-C-G-A-A-G-G-A-G-C-
 A-G-T-G-T-T-T-T-C-G-A-C-C-C-T-T-T-C-T-C-C-C-T-C-G-A-T-G-T-
 50 G-T-G-G-3'

umfaßt oder im wesentlichen homolog damit ist.

- 55 9. Rekombinantes DNA-Plasmid nach einem der vorhergehenden Ansprüche, wobei der genannte Vektor pBR 322 ist.

10. Rekombinantes DNA-Plasmid nach einem der vorhergehenden Ansprüche, das weiters ein Strukturgen

unter der Steuerung dieses Fragments der Pflanzen-DNA enthält, wobei dieses Strukturgen kein natürlicherweise mit diesem Wärmeschockpromotor gefundenes Wärmeschockgen ist.

11. Rekombinantes DNA-Plasmid nach Anspruch 10, wobei das genannte Strukturgen ein Fremdgen ist.
12. Rekombinantes DNA-Plasmid nach Anspruch 11, wobei dieses Fremdstrukturgen ein Herbizidresistenzgen ist.
13. Rekombinantes DNA-Plasmid nach Anspruch 11, wobei dieses Fremdgen ein Z-Gen ist, das für β -Galctosidase codiert.
14. Rekombinantes DNA-Plasmid nach Anspruch 11, wobei dieses Fremdgen ein Gen für kristallines Endotoxin ist.
15. Rekombinantes DNA-Plasmid nach Anspruch 14, wobei dieses Gen für kristallines Endotoxin ein Gen für ein kristallines Endotoxin von Bacillus thuringiensis ist.
16. Rekombinantes DNA-Plasmid nach Anspruch 10, wobei das genannte Strukturgen ein Sojabohnengen ist.
17. Verfahren zur Modifizierung eines Pflanzengenoms, um Expression eines Strukturgens unter der Steuerung eines Sojabohnen-Wärmeschock-Promotorfragmentes zu gestatten, bei welchem Verfahren folgende Schritte vorgesehen sind:
 - (a) Isolierung eines Sojabohnen-Wärmeschock-Genpromotorfragmentes, das die Consensus-Nucleotidsequenz 5'-C-T-X-G-A-A-X-X-T-A-C-X-X-3' umfaßt, worin X für A, T, C oder G steht,
 - (b) Klonierung dieses Sojabohnen-Wärmeschock-Genpromotorfragmentes in einen T-DNA-Shuttle-Vektor zur Gewinnung eines rekombinanten Plasmids,
 - (c) Isolierung eines DNA-Fragmentes, das die Fremdstrukturgene oder die Sojabohnengene, die natürlicherweise nicht unter der regulierenden Steuerung eines Sojabohnen-Wärmeschock-Genpromotors exprimiert werden, trägt, und Einsetzung dieses DNA-Fragmentes in das genannte rekombinante DNA-Plasmid an einer Stelle an der 3'-Seite dieses Sojabohnen-Wärmeschock-Genpromotors, wodurch ein Wärmeschock-Expressionsplasmid gebildet wird, wobei dieses DNA-Fragment im Hinblick auf den genannten Sojabohnen-Wärmeschock-Genpromotor so orientiert ist, daß es unter der Steuerung desselben exprimierbar ist,
 - (d) Transformierung dieses Wärmeschock-Expressionsplasmids in einen ersten Bakterienstamm, der imstande ist, Replikation dieses Wärmeschock-Expressionsplasmids zu unterhalten,
 - (e) Mischen dieses zur Unterhaltung der Replikation dieses Wärmeschock-Expressionsplasmids befähigten Bakterienstammes mit einem zweiten Bakterienstamm, der ein Helferplasmid trägt, das zum Transport dieses Wärmeschock-Expressionsplasmids in einen Agrobacterium-Stamm befähigt ist, der seinerseits die Replikation dieses Wärmeschock-Expressionsplasmids nicht unterhalten kann, wobei dieser Agrobacterium-Stamm ein residentes Plasmid trägt,
 - (f) Selektion auf Rekombination zwischen diesem Wärmeschock-Expressionsplasmid und diesem residenten Plasmid, die ein rekombinantes residentes Plasmid ergibt,
 - (g) Infektion einer Pflanze oder einer Pflanzenzellkultur mit diesem Agrobacterium-Stamm, der dieses rekombinante residente Plasmid enthält und repliziert, und
 - (h) Selektion einer Pflanze oder einer Pflanzenzellkultur, die Pflanzenzellen enthält, die diese Fremdstrukturgene oder diese Sojabohnengene unter der Steuerung dieses aus dem rekombinanten residenten Plasmid in diese Pflanzenzellen übertragenen Sojabohnen-Wärmeschock-Genpromotors enthält, wobei diese Fremdstrukturgene oder diese Sojabohnengene im Anschluß an die Wärmeschockbehandlung oder eine andere Stressbehandlung exprimiert werden.
18. Verfahren nach Anspruch 17, wobei dieses Wärmeschock-Genpromotorfragment eines ist, das in einem der Ansprüche 2 bis 8 genannt wurde.
19. Verfahren nach Anspruch 17 oder Anspruch 18, wobei das in Stufe (c) isolierte DNA-Fragment ein Fremdgen nach einem der Ansprüche 12 bis 15 enthält.
20. Verfahren zur Herstellung einer genetisch modifizierten Pflanzenzelle, bei welchem

- (i) eine Pflanze oder eine Pflanzenzellkultur mit einem Agrobacterium-Stamm infiziert wird, der ein rekombinantes DNA-Plasmid nach einem der Ansprüche 10 bis 16 enthält und repliziert, und
(ii) Pflanzenzellen vermehrt werden, die das genannte Strukturgen unter der Steuerung des genannten Fragmentes der Pflanzen-DNA enthalten.

5

Revendications

1. Plasmide à ADN recombinant comprenant:
 - (a) un vecteur,
 - 10 (b) un fragment d'ADN-T d'un plasmide Ti provenant d'une souche d'*Agrobacterium*; lequel fragment d'ADN-T est fonctionnel pour insertion dans un génome végétal, et
 - (c) un fragment d'ADN végétal qui est capable de réguler l'expression de gènes en réponse à un choc thermique, et lequel fragment est inséré dans ledit fragment d'ADN-T, de manière que ledit fragment d'ADN-T conserve sa fonctionnalité pour insertion dans un génome végétal, ledit fragment
 - 15 d'ADN végétal comprenant un promoteur de choc thermique végétal comprenant la séquence nucléotidique consensus (ou conservée) 5'-C-T-X-G-A-A-X-X-T-A-C-X-X-3', X étant A, T, C ou G.
2. Plasmide à ADN recombinant selon la revendication 1, dans lequel ledit promoteur comprend la séquence nucléotidique consensus 5'-C-T-S-G-A-A-M-R-T-A-C-W-M-K-3', S étant C ou G, M étant A ou C, R étant A ou G, W étant A ou T et K étant T ou G.
- 20 3. Plasmide à ADN recombinant selon la revendication 1, dans lequel ledit promoteur de choc thermique provient du soja.
- 25 4. Plasmide à ADN recombinant selon la revendication 3, dans lequel ledit promoteur de choc thermique est un promoteur de choc thermique de soja choisi parmi hsp2005, hsp2019 et hsp6871.
5. Plasmide à ADN recombinant selon la revendication 1, dans lequel ledit fragment d'ADN végétal est obtenu à partir de la plante soja, le fragment de celui-ci comprenant la séquence

30

5'-T-A-C-A-T-

G-G-T-G-T-G-G-A-G-A-A-T-T-C-A-A-C-C-A-A-A-T-T-G-C-A-A-A-A-
A-G-T-A-G-G-A-T-T-T-T-T-C-T-G-G-A-A-C-A-T-A-C-A-A-G-A-T-T-
35 A-T-C-C-T-T-T-C-A-C-T-T-C-C-T-T-T-A-A-A-T-A-C-C-T-C-G-C-G-
T-A-T-C-C-C-C-T-T-C-G-T-C-C-T-C-G-T-C-A-A-A-C-G-A-A-G-A-A-

40

A-A-A-A-G-T-T-A-C-C-T-G-T-T-T-G-C-G-A-T-C-T-C-A-T-T-A-C-A-
A-T-C-T-C-C-C-T-A-G-T-T-T-C-T-A-A-T-C-T-C-A-G-3',

45

ou étant pratiquement homologue à celle-ci.

6. Plasmide à ADN recombinant selon la revendication 1, dans lequel ledit fragment d'ADN végétal est obtenu à partir de la plante soja, le fragment de celui-ci comprenant la séquence

50

55

5'-T-A-C-A-T-

G-G-T-G-T-G-G-A-G-A-A-T-T-C-A-A-C-C-A-A-A-T-T-G-C-A-A-A-A-

5 A-G-T-A-G-G-A-T-T-T-T-T-C-T-G-G-A-A-C-A-T-A-C-A-A-G-A-T-T-

A-T-C-C-T-T-T-C-A-C-T-T-C-C-T-T-T-A-A-A-T-A-C-C-T-C-G-C-G-

T-A-T-C-C-C-C-T-T-C-G-T-C-C-T-C-G-T-C-A-A-A-C-G-A-A-G-A-A-

10 A-A-A-A-G-T-T-A-C-C-T-G-T-T-T-G-C-G-A-T-C-T-C-A-T-T-A-C-A-

A-T-C-T-C-C-C-T-A-G-T-T-T-C-T-A-A-T-C-T-C-A-G-C-T-A-A-G-A-

A-A-A-A-C-C-A-A-A-A-G-A-T-G-T-C-T-C-T-G-A-T-T-C-C-A-G-G-T-

T-C-C-T-T-C-G-G-T-G-G-C-C-G-A-A-G-G-A-G-C-A-A-C-G-T-C-T-T-

15 C-G-A-T-C-C-A-T-T-C-T-C-A-C-T-C-G-A-C-A-T-G-T-G-G-3',

ou étant pratiquement homologue à celle-ci.

- 20 7. Plasmide à ADN recombinant selon la revendication 1, dans lequel ledit fragment d'ADN végétal est obtenu à partir de la plante soja, le fragment de celui-ci comprenant la séquence

5'-A-G-A-C-C-

25 A-A-T-C-C-T-A-A-C-C-A-A-T-G-T-C-T-G-G-T-T-A-A-G-A-T-G-G-T-

C-C-A-A-T-C-C-C-G-A-A-A-C-T-T-C-T-A-G-T-T-G-C-G-G-T-T-C-G-

A-A-G-A-A-G-C-C-A-G-A-A-T-G-T-T-T-C-T-G-A-A-A-G-T-T-T-C-A-

30 G-A-A-A-A-T-T-C-T-A-G-T-T-T-T-G-A-G-A-T-T-T-T-C-A-G-A-A-G-

T-A-C-G-G-C-A-T-G-A-T-G-A-T-G-C-A-T-A-A-C-A-A-G-G-A-C-T-T-

T-C-T-C-G-A-A-A-G-T-A-C-T-A-T-A-T-T-G-C-T-C-C-T-C-T-A-C-A-

35 T-C-A-T-T-T-T-A-A-A-T-A-C-C-C-C-A-T-G-T-G-T-C-C-T-T-T-G-A-

A-G-A-C-A-C-A-T-C-A-C-A-G-A-A-A-G-A-A-G-T-G-A-A-G-G-C-A-T-

C-G-T-T-A-G-C-A-G-T-T-T-T-G-T-A-G-A-T-T-C-A-A-C-C-T-C-A-A-

T-T-T-G-C-A-G-A-G-T-T-A-C-G-T-T-C-T-A-A-T-A-T-A-T-T-T-A-C-

40 A-C-A-A-G-A-C-T-G-A-C-C-C-3',

ou étant pratiquement homologue à celle-ci.

- 45 8. Plasmide à ADN recombinant selon la revendication 1, dans lequel ledit fragment d'ADN végétal est obtenu à partir de la plante soja, le fragment de celui-ci comprenant la séquence

50

55

5'-A-G-A-C-C-

A-A-T-C-C-T-A-A-C-C-A-A-T-G-T-C-T-G-G-T-T-A-A-G-A-T-G-G-T-
 5 C-C-A-A-T-C-C-C-G-A-A-A-C-T-T-C-T-A-G-T-T-G-C-G-G-T-T-C-G-
 A-A-G-A-A-G-C-C-A-G-A-A-T-G-T-T-T-C-T-G-A-A-A-G-T-T-T-C-A-
 G-A-A-A-A-T-T-C-T-A-G-T-T-T-T-G-A-G-A-T-T-T-T-C-A-G-A-A-G-
 T-A-C-G-G-C-A-T-G-A-T-G-A-T-G-C-A-T-A-A-C-A-A-G-G-A-C-T-T-
 10 T-C-T-C-G-A-A-A-G-T-A-C-T-A-T-A-T-T-G-C-T-C-C-T-C-T-A-C-A-
 T-C-A-T-T-T-T-A-A-A-T-A-C-C-C-C-A-T-G-T-G-T-C-C-T-T-T-G-A-
 A-G-A-C-A-C-A-T-C-A-C-A-G-A-A-A-G-A-A-G-T-G-A-A-G-G-C-A-T-
 15 C-G-T-T-A-G-C-A-G-T-T-T-T-G-T-A-G-A-T-T-C-A-A-C-C-T-C-A-A-
 T-T-T-G-C-A-G-A-G-T-T-A-C-G-T-T-C-T-A-A-T-A-T-A-T-T-T-A-C-
 A-C-A-A-G-A-C-T-G-A-T-A-A-G-A-G-A-A-A-A-T-G-T-C-T-C-T-G-A-
 20 T-T-C-C-A-A-G-T-T-T-C-T-T-C-G-G-T-G-G-C-C-G-A-A-G-G-A-G-C-
 A-G-T-G-T-T-T-T-C-G-A-C-C-C-T-T-T-C-T-C-C-C-T-C-G-A-T-G-T-
 G-T-G-G-3' ,

25 ou étant pratiquement homologue à celle-ci.

9. Plasmide à ADN recombinant selon l'une quelconque des revendications précédentes, dans lequel ledit vecteur est pBR322.
- 30 10. Plasmide à ADN recombinant selon l'une quelconque des revendications précédentes, comprenant en outre un gène structural sous le contrôle dudit fragment d'ADN végétal, lequel gène structural n'est pas un gène de choc thermique rencontré à l'état naturel avec ledit promoteur de choc thermique.
11. Plasmide à ADN recombinant selon la revendication 10, dans lequel ledit gène structural est un gène étranger.
- 35 12. Plasmide à ADN recombinant selon la revendication 11, dans lequel ledit gène structural étranger est un gène de résistance à un herbicide.
13. Plasmide à ADN recombinant selon la revendication 11, dans lequel ledit gène étranger est un gène Z codant pour la β -galactosidase.
- 40 14. Plasmide à ADN recombinant selon la revendication 11, dans lequel ledit gène étranger est un gène d'endotoxine cristalline.
- 45 15. Plasmide à ADN recombinant selon la revendication 14, dans lequel ledit gène d'endotoxine cristalline est un gène d'endotoxine cristalline de *Bacillus thuringiensis*.
16. Plasmide à ADN recombinant selon la revendication 10, dans lequel ledit gène structural est un gène de soja.
- 50 17. Méthode pour la modification d'un génome végétal, de manière à permettre l'expression d'un gène structural sous contrôle d'un fragment promoteur de gène de choc thermique de soja, comprenant les étapes suivantes:
 - 55 (a) isolement d'un fragment promoteur de gène de choc thermique de soja comprenant la séquence nucléotidique consensus 5'-C-T-X-G-A-A-X-X-T-A-C-X-X-3', X étant A, T, C ou G,
 - (b) clonage dudit fragment promoteur de gène de choc thermique de soja dans un vecteur navett à ADN-T, pour l'obtention d'un plasmide à ADN recombinant,

- (c) isolement d'un fragment d'ADN portant des gènes structuraux étrangers ou des gènes de soja non exprimés, naturellement sous le contrôle régulateur d'un promoteur de gène de choc thermique de soja, et insertion dudit fragment d'ADN dans ledit plasmide à ADN recombinant, en une position du côté 3' dudit promoteur de gène de choc thermique de soja, pour l'obtention d'un plasmide d'expression de choc thermique, dans lequel ledit fragment d'ADN est orienté, par rapport audit promoteur de gène de choc thermique de soja, de manière à être exprimable sous contrôle de celui-ci,
- (d) transformation dudit plasmide d'expression de choc thermique dans une première souche bactérienne capable de supporter la réplication dudit plasmide d'expression de choc thermique,
- (e) mélange de ladite souche bactérienne, capable de supporter la réplication dudit plasmide d'expression de choc thermique, avec une seconde souche bactérienne portant un plasmide auxiliaire capable de transporter ledit plasmide d'expression de choc thermique dans une souche d'*Agrobacterium* incapable de supporter la réplication dudit plasmide d'expression de choc thermique, ladite souche d'*Agrobacterium* portant un plasmide résidant,
- (f) sélection pour recombinaison entre ledit plasmide d'expression de choc thermique et ledit plasmide résidant, pour l'obtention d'un plasmide résidant recombinant,
- (g) infection d'une plante ou d'une culture de cellules végétales avec ladite souche d'*Agrobacterium* contenant et répliquant ledit plasmide résidant recombinant, et
- (h) sélection d'une plante ou d'une culture de cellules végétales comprenant des cellules végétales contenant lesdits gènes structuraux étrangers ou lesdits gènes de soja sous contrôle dudit promoteur de gène de choc thermique de soja, transféré dudit plasmide résidant recombinant auxdites cellules végétales, lesdits gènes structuraux étrangers ou lesdits gènes de soja étant exprimés à la suite d'un traitement de choc thermique ou autre traitement de contrainte.
18. Méthode selon la revendication 17, dans lequel ledit fragment promoteur de gène de choc thermique est tel que décrit dans l'une quelconque des revendications 2 à 8.
19. Méthode selon la revendication 17 ou 18, dans lequel le fragment d'ADN isolé dans l'étape (c) comprend un gène étranger selon l'une quelconque des revendications 12 à 15.
20. Méthode pour la production d'une cellule végétale modifiée génétiquement, ladite méthode comprenant
- l'infection d'une plante ou d'une culture de cellules végétales avec une souche d'*Agrobacterium* contenant et répliquant un plasmide à ADN recombinant selon l'une des revendications 10 à 16, et
 - la propagation des cellules végétales contenant ledit gène structural sous contrôle dudit fragment d'ADN végétal.

Figure 1. Restriction sites and sequence of pE2C19

130 140 150 160 170 180
★ ★ ★ ★ ★ ★
ATGATGATGAAAAATGAAAAACCTACTAATGTATTTATGAATAATGTCCAGAAGTGGAA

190 200 210 220 230 240
★ ★ ★ ★ ★ ★
GAAAAATAATATAATGATGTGTAGTAAACAAGAACCTTCGTACATGGTGTGGAGAATTC
MboI RsaI EcoRI
EcoRI
EcoRI*

"CAT" box consensus
250 260 270 280 290 300
★ ★ ★ ★ ★ ★
AACCAAATTGCAAAAAGTAGGATTTTTCTGGAACATACAAGATTATCCTTCACTTCCTT
EcoRI* EcoRI'

"TATA" box
310 320 330 340 350 360
★ ★ ★ ★ ★ ★
TAAATACCTCGCGTATCCCCTTCGTCCTCGTCAAACGAAGAAAAAGTTACCTGTTTGGC
AhaII ThaI MnlI MnlI MboII MboI
Sau3A
+1 transcription

370 380 390 400 410 420
★ ★ ★ ★ ★ ★
ATCTCATTACAATCTCCCTAGTTTCTAATCTCAGCTAAGAAAAACCAAAGATGTCTCTG
DdeIAluI DdeI HinfI

430 440 450 460 470 480
★ ★ ★ ★ ★ ★
ATCCAGGTTTCTTCGGTGGCCGAAGGAGCAACGTCTTCGATCCATTCTCACTCGACATG
BstNI HaeIII MboI MboI MboI
EcoRII Sau3A
MboII TaqI
ScrFI

490 500 510 520 530 540
★ ★ ★ ★ ★ ★
TGGGATCCCTTCAAGGATTTTCATGTTCCCACTTCTTCTGTTTCTGCTGAAAAATTCTGCA
BamHI EcoRI' MboII EcoRI'
MboI EcoRI*
Sau3A

550 560 570 580 590 600
 * * * * * *
 TTCGTGAGCACTCGTGTGGATTGGAAGGAGACCCAGAGGCACACGTGTTCAAGGCTGAT
 HgiAI MnlI

610 620 630 640 650 660
 * * * * * *
 ATTCAGGGGCTGAAGAAAGAGGAAGTCAAGGTTCAAGATTGAAGATGATAGGGTTCTTCAG
 MnlI MboII MboII MboII

670 680 690 700 710 720
 * * * * * *
 ATTAGCGGAGAGAGGAACGTTGAAAAGGAAGACAAGAACGACACGTGGCATCGCGTGGAG
 MnlI MboII ThaISfaNI

730 740 750 760 770 780
 * * * * * *
 CGTAGCAGTGGTAAGTTCACGAGAAGGTTCAAGATTGCCGGAGAATGCAAAAGTGAATGAA
 HpaII

790 800 810 820 830 840
 * * * * * *
 GTGAAGGCTTCTATGGAAAATGGGGTTCTCACTGTCACTGTTCTAAGGAAGAGGTTAAG
 DdeI
 MnlI
 MstII

850 860 870 880 890 900
 * * * * * *
 AAGCCTGATGTTAAGGCCATTGAAATCTCTGGTTGATCCATGTTATGGTTGAAAATCGTG
 MboII HaeIII EcoRI' MboI
 Sau3A

910 920 930 940 950 960
 * * * * * *
 AGCTTATCCTTTGTTGTTGTAATAAGTGTCTTCTGTCTTGTGTGCCTTTGAGAAAAATCT
 AluI MboII MboII EcoRI'

970
 *
 TCCATGCATGCATTGT
 SphI

FIG. 2

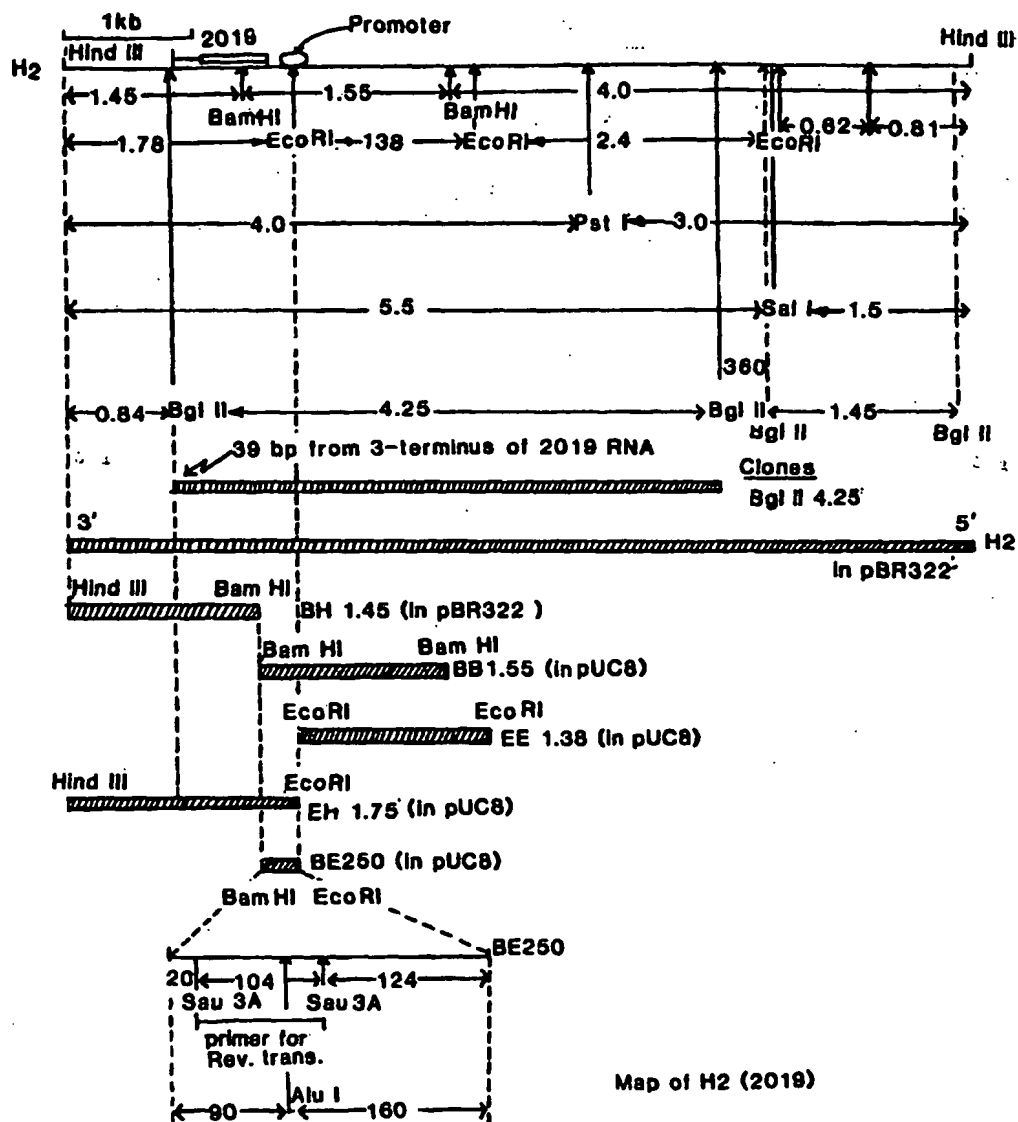


Figure 3.

Nucleotide sequences of pE2019, pM2005, pL2005 and hs6871 heat shock gene promoters

```

pM2005                      AATATTTTGAATAACACATTTT
pL2005                      GAATTTCTGAAATTGGGTCTTTT
hs6871  TTATTTAAAAATACAAATTTTATAAATTAAGTTCAACTCA

pM2005  TTTTAAATATTCTGAAAAATATTTTTCAGAACACAACA
pL2005  TTGTGGGCACTTTTGTATGT-TTTTGTTTAAGTTACTG
hs6871  TCCATATCTCACTCTTTAAATACGATGTTTACTTATTAG

pM2005  ATATTTTCAGAAATTTTATAGGTA-CAAAAGATTTTAAATAA
pL2005  TACTGTGGGGCCACAAAAACGTATAGATCAAAGTAGTAAT
hs6871  ACTCATTAATAAATAAAAAAATAATCATTTGTACAAAG

pE2019                      5'-ATGATGATGAAAAATGGAAAAAC
pM2005  AAAGGATGGTGAATATAGCAAAAGCCTATTTATGAACG
pL2005  AA--TAATATTGATTAAATGATATATATATATATA
hs6871  CCCACCATAAAGGCAATTTGGGCCCTGGTAGACCAATCC

pE2019  CTACTAATGTATTTATGAATAATGTCCAGAAAGTGGAA--
pM2005  ATATCAACCAGAACTAGAAACAAGAAAAATAAATGCACT
pL2005  TATATATATATATCTAG-AAGGTTGTAGAAAGACTAGCT
hs6871  TAACCAATGTCTGCTTAAGATGGTCCAATCCCGAAACT

pE2019  AGAAAAAATAAATATAAT-GAGGTGTAGTAAACAAGA--
pM2005  AGAACCTTCGTACACGGAGTGGAGAAAGTCCAGAAAGTTT
pL2005  AGAACGTACGTA-TTCGTGTGGAGAAAGTCCCTGAAGT--
hs6871  TCTAGTTGCGGTTTCGAAGAAGCCAGAAATGTTTCTGAAA

pE2019  ACCTTCGTACAT-----GGTGTGTGGAGAAATTCAAACCAA
pM2005  TTATAGAATCATTTGAAACTG-----GTAAAAACCAA
pL2005  TTATCGAATCATCTAAAACCTGCTAAAATAGCAAA-CAA
hs6871  GTTTCAGAAAAATTCTAGTTTGTGAGATTTTTCAGAAAGTAC

pE2019  -----ATTGCAAAAAGTAGGATTTTTCTGGAAACATA
pM2005  CCAA--ATTGC-----AACGATTTTTCTGGAAACGTA
pL2005  CATTATATTGT-----CAATATTTTTCTGGAAACATA
hs6871  GGCATGATGATGCATAACAAGGACTTTCTCGAAAGTA
                                     consensus sequence

pE2019  CAAGATTA-TCCTTTCACTTCCTTTAAATAC-----CTCG
pM2005  CACTATTA-TCCTTTCACTTACTTTAAATAC--ATCACG
pL2005  CAAGAGTA-TCCTTTCACTTCCTTTAAATACCTCGAGTG
hs6871  CTATATTGCTCCTCTACATCATTTTAAATACCCCATGTGT

pE2019  CGTATCCCTTCTGTCCTCGTCAAACGAAGA--A-AAA
pM2005  ATTAGTCAGAAAAACGAACGAAGAAGAGAGTTACA-A
pL2005  -TCCCCATTGACTCATAACAAGAGAGAGAGTTACAGA
hs6871  CCTTTGAAGACACATCACAGAAAGAGTGAAGGCATCG

```

Fig. 3 (cont'd.)

```

pE2019 AGTTACCTGTTTGC GATCTCATTACAATCTCCCTAGTT
pM2005 AGTTACCTGTATACGATCTCATTTTGATCTCCCAAGTT
pL2005 A-TTTCCTGTTTACGATCTCATTACAATTTTGCAACTT
hs6871 TTAGCAGTTTGTAGATTCAACCTCAATTTGCAGAGTT

pE2019 TCTAA--TCTCAGCTAAGAAA--A--ACCAAAAGATG
pM2005 TCAA--TCTC-GCGAATAAA-TAT-ATCAAAGATG
pL2005 TCAAAGCTTATTAGC--TAAAGTAACATCAAAGATG
hs6871 ACGTTCTAATATATTTACACAAGACTGATAAGAGAAAATG

```

FIG. 4

Nucleotide sequences of pE2019, pM2005, pL2005 and hs6871 heat shock gene coding sequences

| | |
|--------|---|
| pE2019 | A T G T C T C T G A T T C C A G G T T T C T T C G G T G G C C G A A G G |
| | Met Ser Leu Ile Pro Gly Phe Phe Gly Gly Arg Arg |
| pM2005 | A Ser A T A |
| pL2005 | A T A A T C |
| hs6871 | A Ser Ile Pro |

| | |
|--------|---|
| pE2019 | A G C A A C G T C T T C G A T C C A T T C T C A C T C G A C A T G T G G |
| | Ser Asn Val Phe Asp Pro Phe Ser Leu Asp Met Trp |
| pM2005 | G T C G Val |
| pL2005 | G T |
| hs6871 | G T T C T C T G Val |

| | | | | | | | | | | | | | |
|--------|---|-----|--|-----|-----|-----|-----|-----|------|-----|-----|-----|-----|
| | BamHI | | | | | | | | | | | | |
| pE2019 | G A T C C C T T C A A G G A T T T T C A T G T T C C C A C T T C T T C T | | | | | | | | | | | | |
| | Asp | Pro | | Phe | Lys | Asp | Phe | His | Val | Pro | Thr | Ser | Ser |
| pM2005 | | | | | | | | | T | | | | C T |
| | | | | | | | | | Phe | | | | Leu |
| pL2005 | | | | | | | | | A | | | | |
| | | | | | | | | | Tyr | | | | |
| hs6871 | | C | | | | | | | *Del | | G | | C T |
| | | | | | | | | | | | Ser | | Leu |

| | |
|--------|---|
| pE2019 | G T T T C T G C T G A A A A T T C T G C A T T C G T G A G C A C T C G T |
| | Val Ser Ala Glu Asn Ser Ala Phe Val Ser Thr Arg |
| pM2005 | *Del T A |
| pL2005 | T A A |
| hs6871 | *Del A G T Asn A A |

| | |
|--------|---|
| pE2019 | G T G G A T T G G A A G G A G A C C C C A G A G G C A C A C G T G T T C |
| | Val Asp Trp Lys Glu Thr Pro Glu Ala His Val Phe |
| pM2005 | A A A |
| pL2005 | A C |
| hs6871 | A Gln A |

FIG. 4(cont'd.)

| | |
|--------|---|
| pE2019 | A A G G C T G A T A T T C C A G G G C T G A A G A A A G A G G A A G T C |
| pH2005 | G ^{Lys} Ala Asp Ile Pro Gly Leu Lys Lys G ^{Glu} Glu Val G |
| pL2005 | G |
| hs6871 | G G |
| pE2019 | A A G G T T C A G A T T G A A G A T G A T A G G G T T C T T C A G A T T |
| pH2005 | Lys Val Gln Ile Glu Asp Asp Arg Val Leu Gln Ile A |
| pL2005 | |
| hs6871 | C ^{Leu} G ^{Glu} C ^{Gln} G ^{Gly} C A A |
| pE2019 | A G C G G A G A G A G G A A C G T T G A A A A G G A A G A C A A G A A C |
| pH2005 | Ser Gly Glu Arg Asn Val Glu Lys Glu Asp Lys Asn |
| pL2005 | C ^{Leu} G |
| hs6871 | T A T |
| pE2019 | G A C A C G T G G C A T C G C G T G G A G C G T A G C A G T G G T A A G |
| pH2005 | Asp Thr Trp His Arg Val Glu Arg Ser Ser C ^{Gly} C ^{Lys} C |
| pL2005 | C A Asn |
| hs6871 | T Asp T G |
| pE2019 | T T C A C G A G A A G G T T C A G A T T G C C G G A G A A T G C A A A A |
| pH2005 | Phe Thr Arg Arg Phe Arg Leu Pro Glu Asn Ala Lys |
| pL2005 | T ^{Met} A |
| hs6871 | G ^{Leu} G ^{Val} G T T |

FIG. 4 (cont'd.)

| | | |
|--------|---|--|
| pE2019 | G T G A A T G A A G T G A A G G C T T C T A T G G A A A A T G G G G T T | |
| | Val Asn Glu Val Lys Ala Ser Met Glu Asn Gly Val | |
| pM2005 | G G C | |
| | Glu Gln | |
| pL2005 | C A T *Del G | |
| | Gln Val Cys | |
| hs6871 | G C C C | |
| | Asp Gln | |
| pE2019 | C T C A C T G T C A C T G T T C C T A A G G A A G A G G T T A A G A A G | |
| | Leu Thr Val Thr Val Pro Lys Glu Glu Val Lys Lys | |
| pM2005 | A T C T C A C A G T T T A C C G T G C C C A A A G A A G A | |
| | Leu Thr Val Tyr Arg Ser Gln Arg Arg | |
| pL2005 | T A A A G A A | |
| | Ile Glu | |
| hs6871 | A A | |
| | Ile | |
| pE2019 | C C T G A T G T T A A G G C C A T T G A A A T C T C T G G T T G A T C C | |
| | Pro Asp Val Lys Ala Ile Glu Ile Ser Gly stop | |
| pM2005 | G G T T A A G A A | |
| | Gly stop | |
| pL2005 | A G T C T G A T G T T A A G C C T A T A A G A A | |
| | Ser Leu Met Leu Ser Leu stop | |
| hs6871 | A C A | |
| | Asp | |

Note: Blank spaces indicate identity with the nucleotide sequence or the amino acid sequence.

*Del - indicates a deletion of a codon(s).

FIG. 5

Nucleotide sequences of the regions downstream (i.e., 3'- to the stop codon) from the coding sequences of pE2019, pM2005, pL2005 and hs6871.

```

pE2019  T G A T C C A T G T T A T G G T T G A A A T C G T G A G C T T A T C C T T
pM2005  T A A G A A G G C C T G A T G T T A A G G C C A T T G A A A T C T C T G G T
pL2005  T A A G A A A T C T C T G G T T A A A C T T G G T T T C A C T G A A A A T C
hs6871  T C T A T G T T G C T C T G T T C C T T C G T T G A A A T G T G T T T A T G
      *               *

```

```

pE2019  T G T T G T T G T A A T A A G T G T C T T C T G T C T T G T G T G C C T T T
pM2005  T A A A C T T G T T G T T G C T T A A A A T C G T G A G C T T G C T T G T G
pL2005  G T G A G A G C T T T T A A A T T T G C T T T G T T G T A A T A A G T G T C
hs6871  T T T T C T T A T T C T G A G G A T C A T T T G T G T G A G T C G T G T G A
      *               *

```

```

pE2019  G A G A A A A A T C T T C C A T G C A T G C A T T G T
pM2005  T A T A A C T A A T A A G T A T T C T C G T C A T G T
pL2005  C T T T G T C T T G T G T T C C A A T G G T G A T T T
hs6871  A A A A T A T T T C A G G T T T T A T G T T G G C T A

```

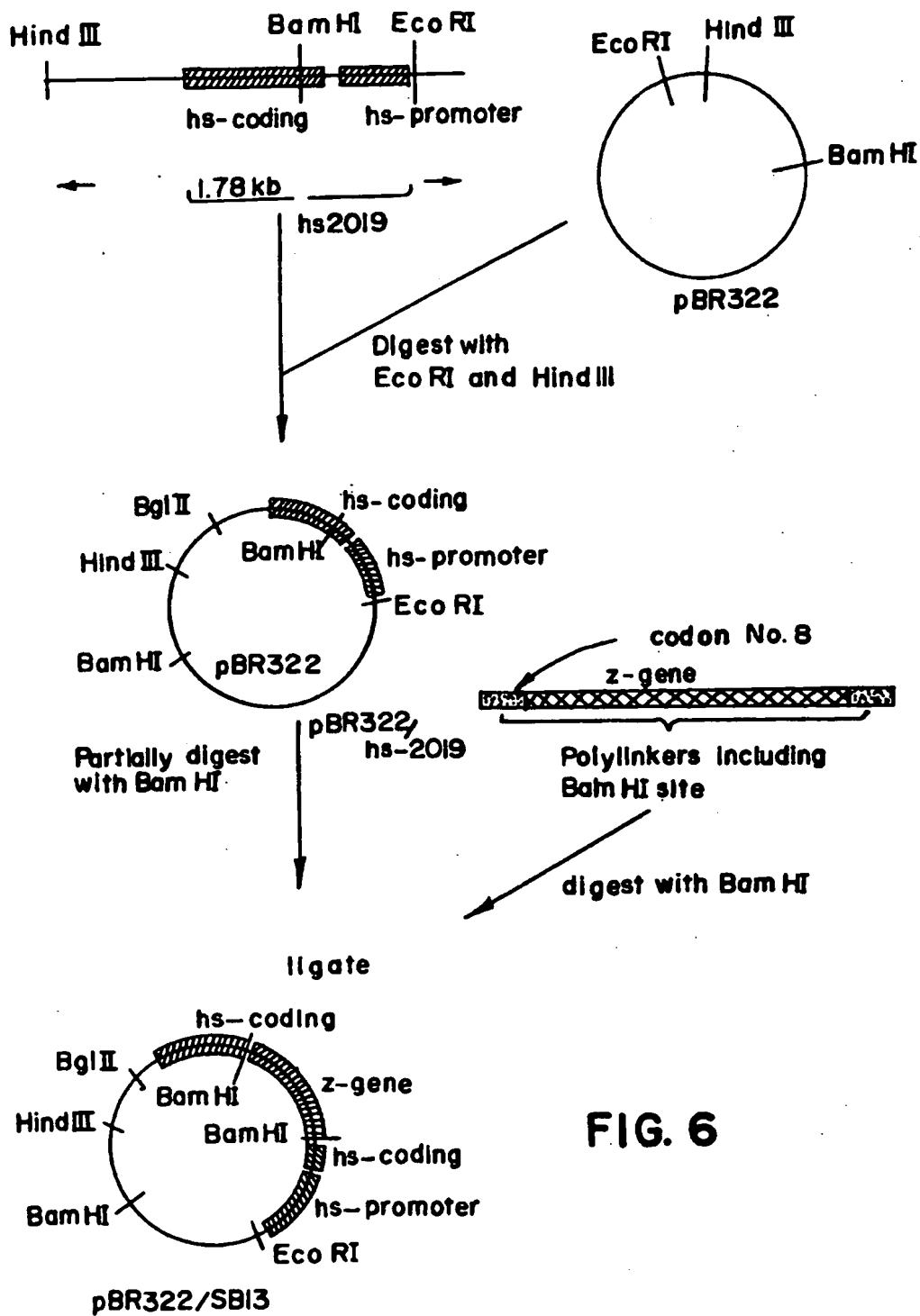


FIG. 6

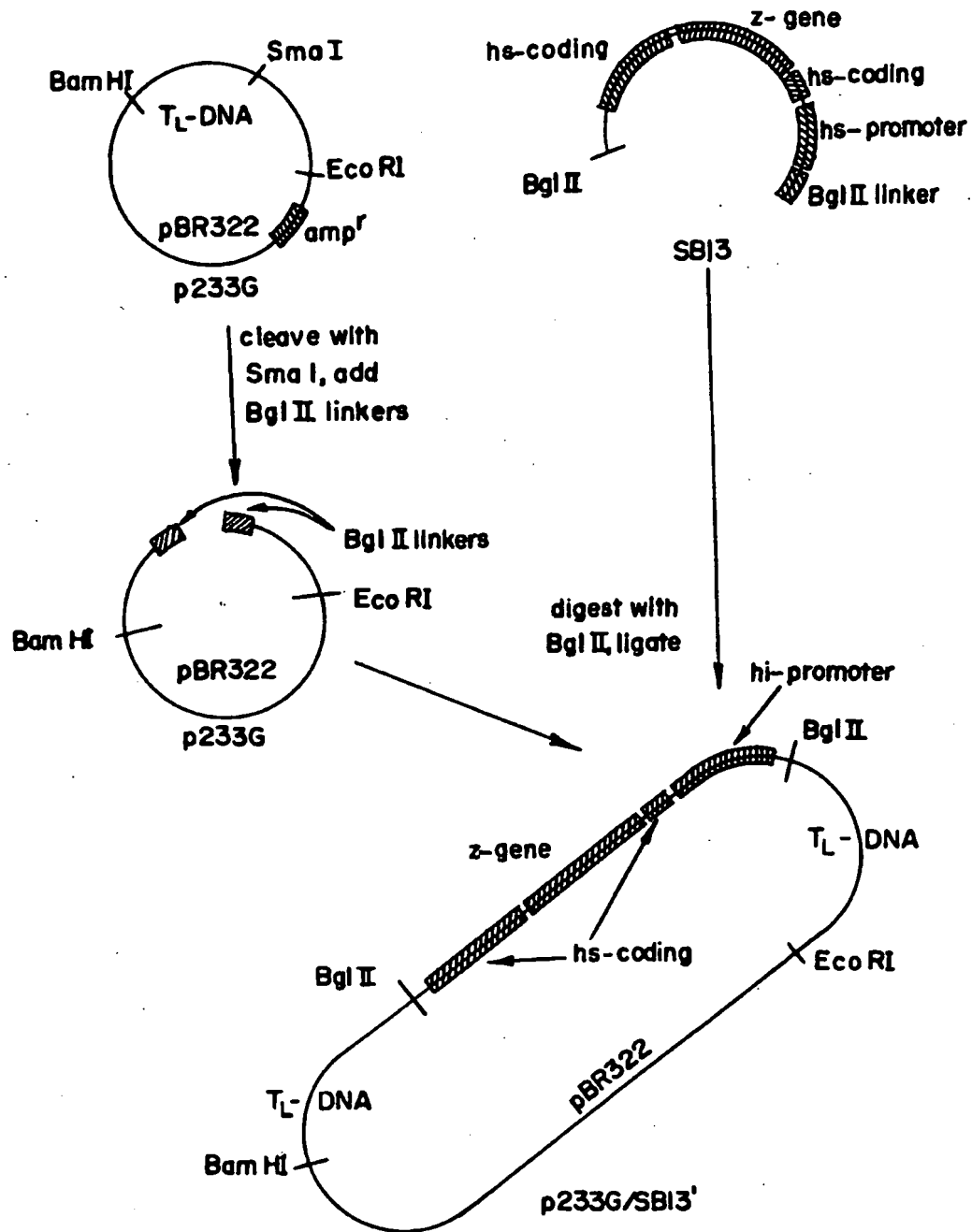


FIG. 7

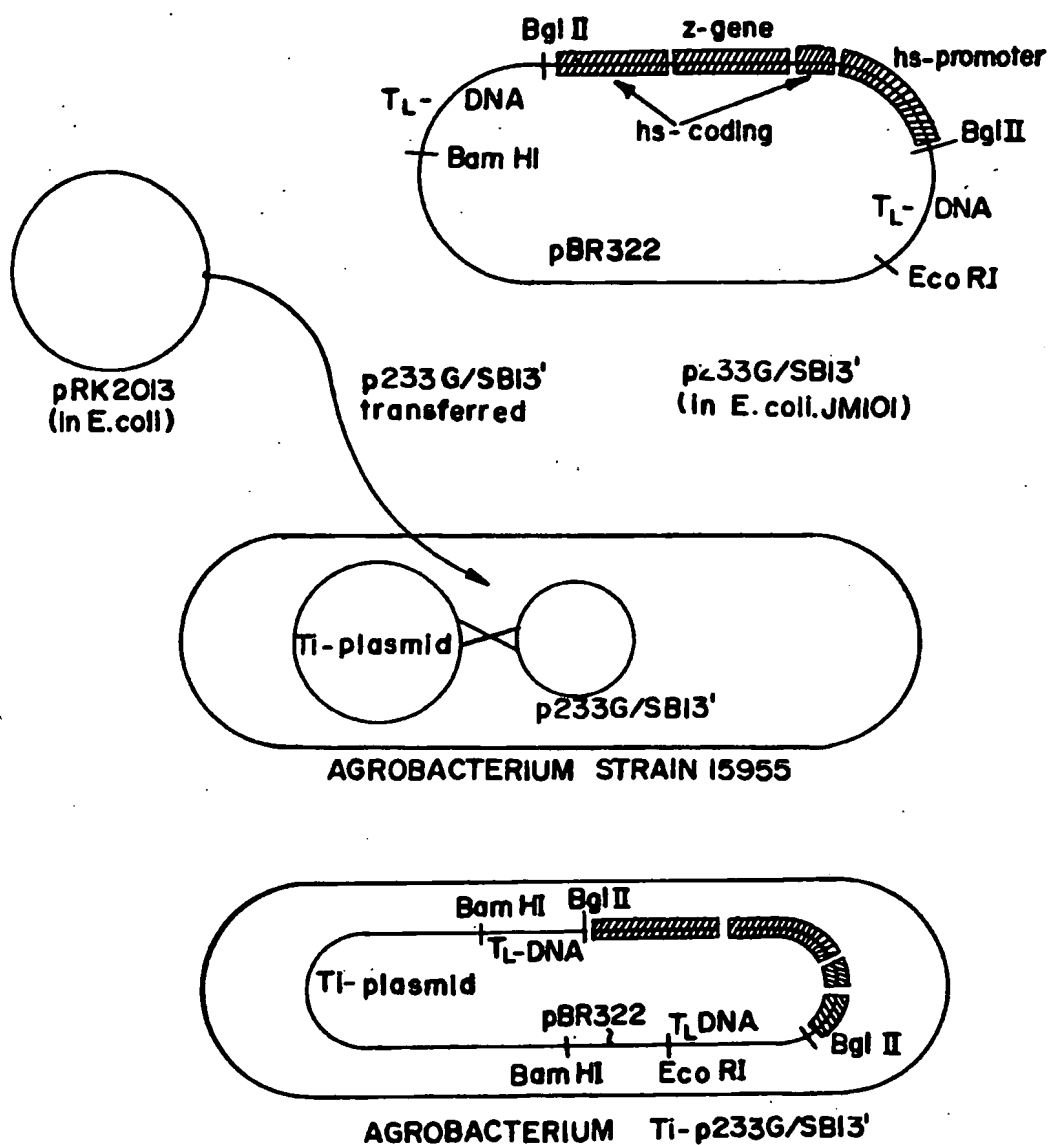


FIG. 8

Restriction sites and sequence of hsp6871

[illegible]

FIG. 9 (cont'd.)

Ile Ser Gly OCURE
 ATC TCT GGT TAA TCTATGTTGCTCTGTCCTGCTGTTGAAATGTTTATGTTTCTTATCTGAGGATCAATTTGTTGTTGAGTCTGTTGAAA +540
 AATATTCAGGTTTATGTTGGCTAAGAGGCTAATGTTTGGGCGCTAGAAAATCTCTGGTTAAACTGTTGTAAGATCTGTTACTTGGTTTAAAGTTTGTG +640
 TGTTTTGTTCACTTCCAGGAATTTATGTTGTCAGGAAGATGTAATTTGAAAAATTTAGCAATAGACTAATGGTTTTATATATCTGTTGTTGCAATAAAT +740
 CTTAGGATATGATATCACTGGAACAGATTCACATATGTCAGTGTGTTGAGAAAGCAATGATAGTTCTTAATCTCCAGTCTACTATGCAATGTTTTTAT +840
 ATTTTAAATTAATATTTTTTATGATGCAATAGAAATTAATGAGACTTTAATAGAAATAGAAATATATACAGTCTCAAGTATGATGATGCAACAGCATCGA 3'

| Restriction Enzyme | Restriction Site | Restriction Enzyme | Restriction Site |
|--------------------|------------------|--------------------|------------------|
| AccI | 278-283 | MboII | 342-346 |
| AhaIII | 313-118 | | 486-490 |
| AsuI | 189-194 | MnlI | 613-617 |
| | 271-275 | | 451-454 |
| | 311-315 | | 540-543 |
| | 665-669 | | 463-468 |
| AsuII | 339-344 | | 653-656 |
| AvaII | 311-315 | RsaI | 244-247 |
| | 665-669 | | 400-403 |
| AvaIII | 413-418 | | 438-441 |
| CfrI | 620-624 | ScaI | 437-442 |
| EcoRII | 274-278 | Ser FI | 274-278 |
| FokI | 174-178 | Sfa NI | 412-416 |
| HaeIII | 272-275 | | 513-517 |
| Hgl AI | 621-624 | TaqI | 1-4 |
| HinfI | 725-730 | | 340-343 |
| | 213-217 | | 432-435 |
| | 533-537 | | 640-643 |
| | 602-606 | XbaI | 655-658 |
| | | | 83-88 |